



UNIVERSIDAD DE BUENOS AIRES.  
Facultad de Ciencias Exactas y Naturales.  
Departamento de Matemática.

Tesis de Licenciatura.

Estrategias de control óptimo para contener un brote de  
Ébola en África Occidental.

Iván Muñoz.

Director: Pablo Amster.

Mayo, 2020.



# Agradecimientos.

Ante todo quisiera agradecerle a Pablo por aceptar dirigir este trabajo, aún sabiendo los desafíos que supondrían trabajar a la distancia. Su guía y observaciones han sido invaluable para mí.

A Willy y Constanza, por aceptar ser jurados de esta tesis. Gracias por su tiempo y valiosas correcciones.

A mis viejos, Maxi y Elsa gracias por todo su apoyo y sacrificio para que yo pueda hacer esta carrera. Sus palabras de aliento en los momentos más difíciles me motivaron a continuar y no desistir.

A mis amigos de la facu, *La Banda de Aplicada*: Joaco, Franco, Naty, Guty, Nico y Nacho. Gracias por su compañía, ayuda, risas y buena onda durante las cursadas, trabajos prácticos y almuerzos. Hicieron mis días en Exactas mucho más felices.

A mi compañera de vida, Caty. Gracias por acompañarme (y aguantarme) durante toda la realización de este trabajo y por ayudarme en la redacción y corrección.



# Introducción.

Durante el 2014 se desarrolló una epidemia de Ébola en algunos países de África occidental. En aquel entonces yo me encontraba estudiando en Buenos Aires y recuerdo que los medios de comunicación cubrían la noticia. Diarios, noticieros y programas de radios hablaban acerca de la letalidad y lo contagioso que era el virus. Mostraban imágenes de médicos y personal de salud cubiertos en trajes protectores donde ni siquiera un centímetro de piel quedaba expuesta. Esto me impresionó. Por suerte, la epidemia pudo ser contenida y no se expandió a nivel mundial.

Nadie se imaginaba que unos años después un virus menos mortal pero más contagioso como el SARS CoV-2 causante del Covid-19 pudiera propagarse tan rápido en casi todos los países del globo y provocar en algunos de ellos el colapso total del sistema de salud. La cuarentena y el aislamiento de toda la población ha sido la principal estrategia de los gobiernos para contener la pandemia. Sin embargo, esto trae aparejado un costo económico: la paralización de la economía. Surgen entonces preguntas como por ejemplo ¿cuánto tiempo se pueden mantener esas medidas?.

En el caso de la epidemia del Ébola, los países afectados se encontraban entre los de menor índice de desarrollo humano a nivel mundial. Sus sistemas de salud eran débiles y no estaban preparados para combatir una enfermedad de ese tipo. Aquí la cuestión era: ¿Cuánta inversión en el sistema de salud sería necesaria para poder contener el brote? ¿Cómo distribuir la inversión a lo largo del tiempo? El denominador común a todas estas preguntas es evitar la expansión de la afección pero tratando de minimizar los costos.

Este trabajo se divide en 4 partes principales. La primera corresponde a una breve recopilación de resultados para el análisis de sistemas dinámicos. En la segunda parte se describirán los modelos matemáticos básicos (y no tan básicos) que se utilizan para explicar la dinámica de las enfermedades infecciosas en una población. También se analizarán sus propiedades matemáticas y resultados fundamentales. La tercera está dedicada a la teoría de control óptimo desde un punto de vista teórico. Finalmente, en la cuarta y última parte, se detallan las características de la enfermedad del virus del Ébola y se propone un modelo para describir el brote del 2014 en África occidental junto con simulaciones numéricas de las estrategias de control óptimo para contenerlo.



# Índice general

|   |           |
|---|-----------|
| <b>1. Preliminares.</b>   | <b>9</b>  |
| 1.1. Sistemas dinámicos. . . . .                                      | 9         |
| 1.2. Sistemas dinámicos planares. . . . .                             | 12        |
| <b>2. Modelos matemáticos de enfermedades infecciosas.</b>            | <b>15</b> |
| 2.1. Breve reseña histórica. . . . .                                  | 15        |
| 2.2. Las formas fundamentales de los modelos de epidemias. . . . .    | 16        |
| 2.3. Modelo SIR. . . . .  | 17        |
| 2.3.1. Propiedades matemáticas del modelo. . . . .                    | 18        |
| 2.3.2. Tasa de contactos adecuados e incidencia. . . . .              | 21        |
| 2.3.3. Tres parámetros importantes: $R_0$ , $\sigma$ , $R$ . . . . .  | 22        |
| 2.4. Dinámica vital. . . . .  | 22        |
| 2.4.1. Propiedades matemáticas del modelo con dinámica vital. . . . . | 23        |
| 2.4.2. Esperanza de vida y edad media de infección. . . . .           | 26        |
| 2.5. Modelos con estrategias de Control. . . . .                      | 27        |
| 2.5.1. Vacunación al nacer. . . . .                                   | 28        |
| 2.5.2. Vacunación continua. . . . .                                   | 30        |
| 2.5.3. Cuarentena y aislamiento. . . . .                              | 32        |
| 2.5.4. Tratamiento. . . . .   | 33        |
| 2.6. Comentarios finales. . . . .                                     | 34        |
| <b>3. Teoría de control óptimo.</b>                                   | <b>37</b> |
| 3.1. Condiciones necesarias. . . . .                                  | 38        |
| 3.1.1. Otras consideraciones. . . . .                                 | 41        |
| 3.1.2. Controles con restricciones. . . . .                           | 44        |
| 3.2. Principio de optimalidad. . . . .                                | 46        |
| 3.2.1. La ecuación de Bellman. . . . .                                | 47        |
| 3.2.2. Interpretación de la adjunta. . . . .                          | 50        |
| 3.3. El Hamiltoniano. . . . .   | 51        |

|   |           |
|---|-----------|
| 3.4. Métodos numéricos . . . . .                                    | 53        |
| 3.4.1. Método de barrido hacia delante y hacia atrás. . . . .       | 55        |
| 3.5. Comentarios finales. . . . .                                   | 56        |
| <b>4. Ébola.</b>  | <b>59</b> |
| 4.1. Características de la enfermedad. . . . .                      | 59        |
| 4.1.1. Signos y síntomas. . . . .                                   | 61        |
| 4.1.2. Transmisión. . . . .   | 62        |
| 4.1.3. Tratamiento. . . . .   | 62        |
| 4.1.4. Prevención y control. . . . .                                | 63        |
| 4.2. Modelando la epidemia. . . . .                                 | 63        |
| 4.2.1. Sistemas de salud en Guinea, Liberia y Sierra Leona. . . . . | 66        |
| 4.2.2. Ajustando los parámetros. . . . .                            | 67        |
| 4.2.3. Definiendo el control y evaluando los costos. . . . .        | 71        |
| 4.3. Resultados obtenidos. . . . .                                  | 74        |
| 4.4. Discusión de los resultados. . . . .                           | 80        |
| 4.5. Comentarios finales. . . . .                                   | 82        |
| <b>A. Métodos numéricos.</b>  | <b>83</b> |
| A.1. Reglas de cuadraturas. . . . .                                 | 83        |
| A.1.1. Reglas de cuadraturas compuestas. . . . .                    | 84        |
| A.2. Métodos de Runge-Kutta. . . . .                                | 85        |
| <b>Bibliografía</b>   | <b>87</b> |



# Capítulo 1

## Preliminares.

El objetivo de este capítulo es presentar las herramientas matemáticas básicas para el estudio de los sistemas dinámicos. Se brindan de manera concisa las definiciones necesarias para enunciar las proposiciones y teoremas que permitan establecer propiedades fundamentales de los sistemas en cuestión.

Lo que sigue a continuación está basado en [40].

### 1.1. Sistemas dinámicos.

Sea  $M \subseteq \mathbb{R}^n$  un conjunto abierto y una función  $F : M \rightarrow \mathbb{R}^n$ . Supongamos que  $F$  es lo suficientemente buena (por ej.  $F \in C^k(M, \mathbb{R}^n)$ ,  $k \geq 1$ ) para que exista una única solución  $X : I \subseteq \mathbb{R} \rightarrow M$  del problema de valores iniciales o IVP:

$$\begin{aligned} X' &= F(X) \\ X(0) &= X_0 \end{aligned} \tag{1.1}$$

Ahora bien, a medida que variamos el dato inicial  $X_0 \in M$  varía la solución del sistema. En particular, las soluciones de (1.1) se denominan curvas integrales o *trayectorias*. Diremos que  $\phi$  es una trayectoria en  $X_0$  si es solución del IVP y además  $\phi(0) = X_0$ . Dada la bondad de  $F$  existe una única trayectoria maximal  $\phi_X$  en  $X \in M$  definida en un intervalo maximal  $I_X = (T_-(X), T_+(X))$ . Introducimos el conjunto

$$W = \bigcup_{X \in M} I_X \times \{X\} \subseteq \mathbb{R} \times M$$

y definimos el flujo de la ecuación diferencial al operador

$$\begin{aligned} \Phi : W &\rightarrow M \\ (t, X) &\mapsto \phi(t, X) \end{aligned}$$

donde  $\phi(t, X)$  es la trayectoria maximal en  $X$ . A veces se usará la notación  $\Phi_X(t) = \Phi(t, X)$  y  $\Phi_t(X) = \Phi(t, X)$ .

Notar que valen las siguientes propiedades:

1.  $\Phi_0 : M \rightarrow \mathbb{R}^n$  es la identidad:  $\Phi_0(X) = X \forall X \in M$ .
2. Si  $\phi(\cdot)$  es la trayectoria maximal en  $X$ , entonces  $\phi(\cdot + s)$  es la trayectoria maximal en  $Y = \phi(s)$  y en particular  $I_X = s + I_Y$ . Como consecuencia, para cada  $X \in M$  y  $s \in I_X$  se tiene que

$$\Phi(s + t, X) = \Phi(t, \Phi(s, X))$$

para todo  $t \in I_{\Phi(s, X)} = I_X - s$ .

En este trabajo llamaremos *sistema dinámico* al sistema de ecuaciones diferenciales autónomo y su flujo asociado.

**Definición 1.1.1.** La *órbita* de  $X$  está definida por

$$\gamma(X) = \Phi(I_X \times X) \subseteq M.$$

Notar que  $Y \in \gamma(X)$  implica que  $Y = \Phi(t, X)$  para algún  $t \in I_X$ , y por lo tanto se tiene que  $\gamma(Y) = \gamma(X)$ . En particular, órbitas distintas son disjuntas. Si  $\gamma(X) = \{X\}$ , entonces  $X$  se denomina *punto fijo* de  $\Phi$  (otras variantes son punto singular, estacionario o de equilibrio). Si lo anterior no pasa entonces se dice que  $X$  es *un punto regular* y  $\Phi(\cdot, X) : I_X \hookrightarrow M$  es inyectiva (o periódica).

Similarmente, se definen las *órbitas positivas y negativas* como

$$\gamma_{\pm}(X) = \Phi((0, T_{\pm}(X)), X).$$

Obviamente  $\gamma(X) = \gamma_-(X) \cup \{X\} \cup \gamma_+(X)$ . Se dice que  $X$  es un *punto periódico* de  $\Phi$  si existe un  $T > 0$  tal que  $\Phi(T, X) = X$ . La menor de las cotas inferiores de tales  $T$  se denominará *período*, es decir,  $T(X) = \inf \{T > 0 \mid \Phi(T, X) = X\}$ . Gracias a la continuidad de  $\Phi$  vale que  $\Phi(T(X), X) = X$  y por la propiedad del flujo  $\Phi(t + T(X), X) = \Phi(t, X)$ . Diremos que una órbita es periódica si un (y por lo tanto todos) punto de la misma es periódico.

No es difícil ver que  $X$  es periódico si y sólo si  $\gamma_+(X) \cap \gamma_-(X) \neq \emptyset$  y por lo tanto, a veces las órbitas periódicas serán también llamadas *órbitas cerradas*.

Se podrían clasificar entonces las órbitas de  $F$  de la siguiente forma:

- *Órbitas fijas* (correspondientes a puntos periódicos con período cero).
- *Órbitas periódicas regulares* (correspondientes a puntos periódicos con períodos positivos).
- *Órbitas no cerradas* (correspondientes a puntos no periódicos).

La cantidad  $T_+(X) = \sup I_X$  (respectivamente  $T_-(X) = \inf I_X$ ) definida antes es llamada *tiempo de vida* positivo (respectivamente negativo) de  $X$ . Un punto  $X \in M$  se dice  *$\sigma$  completo*,  $\sigma \in \{+, -\}$ , si  $T_{\sigma}(X) = \sigma\infty$  y *completo a secas* si es  $+$  completo y  $-$  completo, es decir,  $I_X = \mathbb{R}$ .

**Lema 1.1.1.** Sea  $X \in M$  y supongamos que la órbita positiva (respectivamente negativa) está contenida en un compacto  $C$  de  $M$ . Entonces  $X$  es  $+$  (respectivamente  $-$ ) completo.

**Definición 1.1.2.** Si todos los puntos son completos, entonces el campo vector se dice *completo*. Por lo tanto, si  $F$  es completo significa que  $\Phi$  está globalmente definido, es decir,  $W = \mathbb{R} \times M$ .

**Definición 1.1.3.** Un conjunto  $U \subseteq M$  se dice positivamente invariante si

$$\gamma_+(X) \subseteq U \quad \forall X \in U.$$

Análogamente para negativamente invariante. Un conjunto  $U \subseteq M$  se denomina invariante a secas si es positiva y negativamente invariante.

**Definición 1.1.4.** El conjunto  $\omega_{\pm}$ -límite de un punto  $X \in M$ ,  $\omega_{\pm}(X)$ , es el conjunto de aquellos puntos  $Y \in M$  para los cuales existe una sucesión  $t_n$  tal que  $t_n \rightarrow \pm\infty$  con  $\Phi(t_n, X) \rightarrow Y$ .

*Notación:* De aquí en adelante utilizaremos  $\sigma \in \{+, -\}$ .

**Lema 1.1.2.** El conjunto  $\omega_{\sigma}(X)$  es cerrado e invariante.

**Lema 1.1.3.** Si  $\gamma_{\sigma}(X)$  está contenido en un compacto, entonces  $\omega_{\sigma}(X)$  es no vacío, compacto y conexo.

**Lema 1.1.4.** Si  $\gamma_{\sigma}(X)$  está contenido en un compacto, entonces

$$\lim_{t \rightarrow \sigma\infty} d(\Phi(t, X), \omega(X)) = 0.$$

**Definición 1.1.5.** Un punto fijo  $X_0$  de  $F(X)$  es llamado (Liapunov) *estable* si para cualquier vecindad  $U(X_0)$  existe otra vecindad  $V(X_0) \subseteq U(X_0)$  tal que cualquier solución empezando en  $V(X_0)$  permanece en  $U(X_0)$  para todo  $t \geq 0$ . Un punto que no es estable se llamará *inestable*. Similarmente un punto fijo  $X_0$  de  $F(X)$  se denominará *asintóticamente estable* si es estable y existe una vecindad  $U(X_0)$  tal que

$$\lim_{t \rightarrow \infty} |\Phi(t, X) - X_0| = 0 \quad \text{para todo } X \in U(X_0).$$

Finalmente un punto fijo  $X_0$  de  $F(X)$  es *exponencialmente estable* si existen constantes  $\delta, \alpha, C > 0$  tales que

$$|\Phi(t, X) - X_0| \leq Ce^{-\alpha t}|X - X_0| \quad \text{para todo } X \in B_{\delta}(X_0).$$

Notar que un punto exponencialmente estable también es estable y asintóticamente estable.

**Teorema 1.1.1.** (Estabilidad exponencial via linealización) Supongamos que  $F \in C^1$  tiene un punto fijo  $X_0$  y que todos los autovalores de la matriz Jacobiana evaluada en  $X_0$  tienen parte real negativa. Entonces  $X_0$  es exponencialmente estable.

**Definición 1.1.6.** Sea  $X_0$  un punto fijo de  $F(X)$  y sea  $U(X_0)$  una vecindad abierta de  $X_0$ . Una *función de Liapunov* es una función continua

$$L : U(X_0) \rightarrow \mathbb{R}$$

la cual cumple que vale cero en  $X_0$ , es positiva para  $X \neq X_0$  y satisface

$$L(\phi(t_0)) \geq L(\phi(t_1)) \text{ con } t_0 < t_1 \text{ y } \phi(t_0), \phi(t_1) \in U(X_0) \setminus \{X_0\} \quad (1.2)$$

para cualquier solución  $\phi(t)$ . La  $L$  de arriba se dice *una función estrictamente de Liapunov* si la igualdad en (1.2) nunca ocurre.

**Teorema 1.1.2.** (Liapunov) Supongamos que  $X_0$  es un punto fijo de  $F$ . Si existe una función de Liapunov  $L$  entonces  $X_0$  es estable.

**Teorema 1.1.3.** (Principio de Krasovskii - LaSalle) Supongamos que  $X_0$  es un punto fijo de  $F$ . Si existe una función de Liapunov  $L$  que no es constante en cualquier órbita totalmente contenida en  $U(X_0) \setminus \{X_0\}$ , entonces  $X_0$  es asintóticamente estable. Más aún, cualquier órbita totalmente contenida en  $U(X_0)$  converge asintóticamente a  $X_0$ .

## 1.2. Sistemas dinámicos planares.

**Proposición 1.2.1.** Supongamos que  $\omega_\sigma(X) \cap \gamma_\sigma(X) \neq \emptyset$ . Entonces  $X$  es periódico y se tiene que  $\omega_+(X) = \gamma(X) = \omega_-(X)$ .

**Proposición 1.2.2.** Sea  $C$  un conjunto  $\sigma$  invariante minimal. Entonces  $C$  es una órbita periódica.

**Teorema 1.2.1.** (Teorema de Poincaré-Bendixson) Si  $w_\sigma(X) \neq \emptyset$  es compacto y no contiene puntos de equilibrio, entonces  $w_\sigma(X)$  es una órbita cerrada regular.

**Proposición 1.2.3.** Supongamos que  $\omega_\sigma(X)$  contiene a una órbita periódica regular  $\gamma(Y)$ . Entonces  $\omega_\sigma(X) = \gamma(Y)$ .

**Definición 1.2.1.** Definimos los siguientes conceptos:

1. Sea  $X_0$  un punto fijo y silla de  $F$ . Una órbita  $\gamma(Y)$  se dice *homóclina* si  $\omega_-(Y) = \omega_+(Y) = \{X_0\}$  ( $\gamma(Y)$  conecta un punto silla con sí mismo).
2. Sea  $X_0, X_1$  dos puntos fijos distintos de  $F$ . Una órbita  $\gamma(Y)$  se dice *heteróclina* si  $\omega_-(Y) = \{X_0\}$  y  $\omega_+(Y) = \{X_1\}$  ( $\gamma(Y)$  conecta dos puntos de equilibrio distintos).
3. Informalmente digamos que una *variedad  $n$ -dimensional* es un espacio matemático que a pequeña escala se parece a un espacio Euclídeo de  $n$  dimensiones. Por ejemplo, una línea y un círculo son variedades 1-dimensionales, mientras que un plano o una esfera son variedades 2-dimensionales.

4. Una *separatrix* es una órbita en el espacio de fases que une un equilibrio hiperbólico con sí mismo o que conecta las variedades estables e inestables de un par de puntos de equilibrio distintos. Una separatrix marca la frontera entre sectores de espacios de fase con distintas propiedades.
5. Un *ciclo separatrix* consiste en la unión de finitos puntos de equilibrio  $\{X_j\}_{1 \leq j \leq J}$  y separatrices  $\{\gamma_j\}_{1 \leq j \leq J}$  tales que el flujo en  $\gamma_j$  es desde  $X_j$  hacia  $X_{j+1}$  y  $X_{J+1} = X_1$ .
6. Un *ciclo separatrix compuesto* o *gráfico* es la unión de finitos ciclos separatrices compatiblemente orientados.

**Teorema 1.2.2.** (Teorema de Poincaré-Bendixson generalizado) Sea  $M$  un subconjunto abierto de  $\mathbb{R}^2$  y sea  $F \in C^1(M, \mathbb{R}^2)$ . Fijemos  $X \in M$ ,  $\sigma \in \{+, -\}$ , y supongamos que  $\omega_\sigma(X) \neq \emptyset$  es compacto, conexo y contiene a lo sumo una cantidad finita de puntos fijos. Entonces se da sólo uno de los siguientes casos:

1.  $\omega_\sigma(X)$  es una órbita fija.
2.  $\omega_\sigma(X)$  es una órbita periódica regular.
3.  $\omega_\sigma(X)$  es un ciclo separatrix compuesto o gráfico. ( $\omega_\sigma(X)$  consiste en finitos puntos de equilibrio  $\{X_j\}$  y órbitas no-cerradas  $\gamma(Y)$  tales que  $\omega_\pm(Y) \in \{X_j\}$ .)

**Proposición 1.2.4.** El interior de toda órbita cerrada debe contener un punto de equilibrio.

**Teorema 1.2.3.** (Criterio de Bendixson) Supongamos que la divergencia de  $F$ ,  $\text{div}(F)$ , no cambia de signo ni se anula idénticamente en una región simplemente conexa  $U \subseteq M$ . Entonces no hay órbitas regulares periódicas ni gráficos (totalmente) contenidos en  $U$ .

**Teorema 1.2.4.** (Criterio de Dulac) Supongamos que existe una función escalar  $\alpha(X)$  tal que  $\text{div}(\alpha F)$  no cambia de signo ni se anula idénticamente en una región simplemente conexa  $U \subseteq M$ . Entonces no hay órbitas regulares periódicas ni gráficos (totalmente) contenidos en  $U$ .

**Proposición 1.2.5.** Si la intersección  $\omega_+(X) \cap \omega_-(X) \neq \emptyset$  contiene un punto regular, entonces  $X$  es periódico.



## Capítulo 2

# Modelos matemáticos de enfermedades infecciosas.

El objetivo de este capítulo es comentar brevemente la historia de las enfermedades infecciosas y cómo estas afectaron a la humanidad. Los científicos a lo largo del tiempo eligieron distintas técnicas para estudiar la propagación de epidemias en las poblaciones con el fin de brindar mayor entendimiento acerca de factores fundamentales que dominan la dinámica, y de esta forma poder desarrollar estrategias para su erradicación. Es así como se llegó a los modelos matemáticos epidemiológicos. Más adelante, se describirán varios de los modelos fundamentales, presentando también las herramientas matemáticas necesarias para su estudio.

### 2.1. Breve reseña histórica.

Las enfermedades infecciosas han azotado a la humanidad desde su comienzo. A lo largo de la historia estas han causado desde crisis sociales y económicas hasta contribuido a la caída de imperios. Instamos al lector interesado que recurra a [56] para conocer las grandes epidemias que han afectado al mundo a lo largo del tiempo. La Plaga de Antonina 165 - 180 dc, fue una antigua pandemia traída al Imperio Romano por las tropas que retornaban de las campañas de Este. La verdadera causa se mantiene incierta aunque se sospecha que haya sido causada por el virus de la viruela o el sarampión. Se estima que la plaga se cobró las vidas de un tercio de población en algunas áreas, entre ellas la del emperador Lucius Verus y devastó al ejército romano. [57]

La Peste Negra afectó gravemente a Europa durante el siglo XIV y era transmitida por unas pulgas transportadas por ratas. La información sobre la mortalidad varía ampliamente entre las fuentes, pero se estima que entre el 30 % y el 60 % de la población de Europa murió desde el comienzo del brote a mitad del siglo XIV. Esto trajo como consecuencia una gran movilidad de los pocos individuos sobrevivientes, causando a veces la despoblación entera de ciertas localidades. Dichos movimientos, al mismo tiempo, alimentaban la pandemia ya que se transportaba la enfermedad de un lugar a otro. [53]

La Gran Peste (1665-1666) fue una epidemia que mató entre 70 000 y 100 000 personas en Inglaterra, y más de una quinta parte de la población de Londres. La poca higiene del momento fue sin duda alguna un factor determinante en esta epidemia, como en muchas otras de la época. Esta en particular obligó al rey y su corte a trasladarse a la ciudad de Oxford, lugar en el que aparentemente estarían a salvo. [55]

En la era moderna, enfermedades como el VIH/SIDA afectan millones de personas en todo el mundo. En 2006, se ha calculado que 39.5 millones de personas han tenido el VIH/SIDA. De ellas, casi tres millones han muerto [50, 51]. La epidemia de Ébola de 2014-2016 fue el mayor brote epidémico de la enfermedad por el virus del Ébola, originado en diciembre de 2013 en Guinea. El virus del Ébola causa en el ser humano la enfermedad del Ébola, cuya tasa de letalidad puede llegar al 90%. [54]

Estas enfermedades han puesto en jaque a cada uno de los gobiernos de la época como así también a los científicos y organizaciones encargadas de combatirlas. En la actualidad el uso de herramientas como la computación, simulaciones numéricas y modelos epidemiológicos han brindado una comprensión global más extensa de la dinámica de las enfermedades infecciosas. Esto permite proveer predicciones y guías útiles para que mejores estrategias sean establecidas.

## **2.2. Las formas fundamentales de los modelos de epidemias.**

Estos modelos fueron estudiados y propuestos originalmente por Kermack y McKendrick (1927, 1932). Están basados en la división en compartimientos o clases de una población. Son modelos determinísticos de ecuaciones diferenciales que describen la dinámica de una enfermedad dentro de la población. Hoy en día, se ha avanzado drásticamente en la complejidad de los mismos y se han convertido en importantes herramientas para analizar la propagación y el control de las enfermedades infecciosas.

El entendimiento de las características de transmisión de las enfermedades en las comunidades, regiones y países lleva a mejores aproximaciones en la disminución del contagio de estas afecciones. Los modelos matemáticos se usan para comparar, planear, implementar, evaluar y optimizar los programas de detección, prevención, terapia y control.

En muchas ciencias es posible llevar a cabo experimentos para obtener información y testear hipótesis. Sin embargo, los experimentos relacionados con la propagación de enfermedades infecciosas son generalmente imposibles, no éticos o demasiado costosos. A veces la información está disponible debido a epidemias que ocurren naturalmente y por la incidencia natural de las enfermedades endémicas. No obstante, los datos suelen estar incompletos debido a la falta de informes. Esta falta de información hace difícil la estimación de los parámetros, por lo que a veces, sólo es posible estimar un rango de valores. Una forma de compensar la falta de experimentos o datos precisos, es llevar adelante simulaciones computacionales que repliquen la dinámica para un intervalo de valores o conjuntos.



Los modelos matemáticos tienen tanto limitaciones como capacidades que deben ser reconocidas. Algunas preguntas no pueden ser contestadas a través de ellos, pero cada tanto un modelador es capaz de encontrar la combinación adecuada de información disponible, una pregunta interesante y un modelo matemático que lo conduzca a una respuesta [19].

### 2.3. Modelo SIR.

En este modelo la población que se estudia es dividida en compartimientos o clases disjuntas cuyos tamaños varían en el tiempo. Se asume que la cantidad total de individuos se mantiene constante,  $N$ , y que es lo suficientemente grande para que el tamaño de cada clase sea considerado como una variable continua. Adicionalmente se asume que la población está homogéneamente mezclada y que la distribución espacial de los individuos no es un factor que afecte a la dinámica de la enfermedad. Los compartimientos se diferencian por el estado de los individuos en base a la enfermedad. La primera clase es la de los susceptibles, que denotaremos  $S$ , y alberga a dichos individuos que no se encuentran infectados por la enfermedad pero tampoco la han padecido en el pasado. Estos pueden contraerla si se exponen a la misma. La segunda clase será la de los infectados, que llamaremos  $I$ , y contiene a aquellos que se encuentran infectados y padecen la enfermedad (infectiosos). La última de las clases, es la de los removidos o recuperados. Esta se caracteriza por tener a aquellos que han sufrido la enfermedad y generado inmunidad permanente. Denotaremos a la cantidad de individuos en las clases  $S, I$ , y  $R$  a tiempo  $t$  como  $S(t), I(t)$  y  $R(t)$  respectivamente. Se propone entonces el siguiente sistema de ecuaciones diferenciales ordinarias para modelar la dinámica de la enfermedad en la población:

$$\begin{aligned}\frac{dS}{dt} &= S'(t) = -\beta S(t)I(t) \\ \frac{dI}{dt} &= I'(t) = \beta S(t)I(t) - \gamma I(t) \\ \frac{dR}{dt} &= R'(t) = \gamma I(t)\end{aligned}\tag{2.1}$$

Adicionalmente se han hecho los siguientes supuestos: La población total se encuentra aislada, es decir no hay migraciones de ningún tipo. Tampoco se consideran nacimientos o muertes. De esta manera como se ha observado antes su número total se mantiene constante,  $S(t) + I(t) + R(t) \equiv N$ . El tiempo de incubación es tan corto que se considera despreciable; esto es, un individuo susceptible que contrae la enfermedad es infeccioso al instante.  $\beta$  se denomina *coeficiente de transmisión* y  $\gamma$  representa el *coeficiente de la tasa de recuperación*. Se asume que la cantidad de agentes susceptibles que son infectados por un agente infeccioso por unidad de tiempo, a tiempo  $t$ , es proporcional a la cantidad de susceptibles multiplicado por el coeficiente de transmisión, *incidencia de acción de masa*. Por último,  $\gamma I(t)$  representa a la cantidad de individuos infecciosos que se recuperan por unidad de tiempo, a tiempo  $t$ .

Retomando la ecuación (2.1) para que el problema esté bien planteado matemáticamente, es necesario definir condiciones iniciales,  $S(0) = S_0, I(0) = I_0, R(0) = 0$  (pues consideramos que la población inicial de recuperados es nula, es decir, ningún agente posee inmunidad a la infección). Recordemos que la población se mantiene constante en un valor  $N$  a lo largo del tiempo, por lo tanto debe suceder que  $S_0 + I_0 = N$ . Además dado que estos valores representan cantidades de individuos deben ser no negativos, es decir,  $S_0, I_0 \geq 0$ . Observemos también que esto debe prolongarse en el tiempo,  $S(t), I(t), R(t) \geq 0 \forall t$ .

Este tipo de sistemas dinámicos pueden ser esquemáticamente representados a través de diagramas de flujo donde cada compartimiento es simbolizado con una caja indexada con el nombre de la clase. Las flechas indican la dirección de movimiento de los individuos a través de los compartimientos y sobre éstas se muestra la variación por unidad de tiempo.

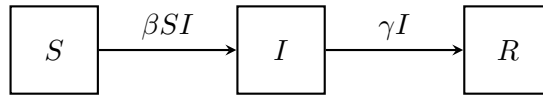


Figura 2.1: Diagrama de flujo, SIR.

### 2.3.1. Propiedades matemáticas del modelo.

Primeramente recordemos la ecuación (2.1) y notemos lo siguiente:

1. Dado que  $S'(t) \leq 0 \forall t$  el número de susceptibles es no creciente independientemente de la condición inicial  $S_0$ . Recordemos también que debido a que la población se mantiene constante en un valor  $N$  y  $S(t)$  es no negativo, tenemos entonces que  $S(t)$  es acotado y no creciente para todo  $t$ . Podemos concluir entonces que existe el límite cuando  $t$  tiende a infinito.

$$\lim_{t \rightarrow \infty} S(t) = S_\infty$$

2. Similarmente notemos que  $R(t)$  también posee un comportamiento monótono pero no decreciente, ya que  $R'(t) \geq 0 \forall t$ . De la misma forma que  $S$ , sabemos que  $R(t)$  está acotado por 0 y  $N$ . Concluimos entonces que existe el límite.

$$\lim_{t \rightarrow \infty} R(t) = R_\infty$$

3. Por otro lado, el comportamiento de los infecciosos es diferente. Su dinámica viene dada por  $I'(t) = I(t)(\beta S(t) - \gamma)$ . Notar entonces que el signo de  $I'(t)$  es igual al signo de  $\beta S(t) - \gamma$  ya que  $I(t)$  es siempre no negativo. Puede pasar que los infectados comiencen y se mantengan decreciendo o que crezcan al principio y

luego decrezcan (en este caso se habla de una epidemia o de un brote)<sup>1</sup>. Para esto último es condición necesaria y suficiente que  $\beta S(0) - \gamma > 0$ , o lo que es lo mismo que  $\frac{\beta}{\gamma} S(0) > 1$ . Cualquiera sea el caso, gracias a que  $S(t) + I(t) + R(t) \equiv N \forall t$ , se puede afirmar que también existe el límite para la clase de los infectados. Digamos que

$$\lim_{t \rightarrow \infty} I(t) = I_{\infty}.$$

4. Se puede mostrar además que  $S_{\infty} > 0$ . Para esto en el sistema (2.1) dividamos la primera ecuación por la tercera obteniendo

$$\frac{S'}{R'} = -\frac{\beta}{\gamma} S.$$

Resolviendo llegamos a que

$$S(t) = S(0)e^{-\frac{\beta}{\gamma} R(t)} \geq S(0)e^{-\frac{\beta}{\gamma} N} = \text{Cte} > 0.$$

5. Más aún,  $I_{\infty} = 0$ . Luego, se puede afirmar que la enfermedad siempre se extingue. La razón principal de esto resulta ser la falta de infecciosos y no la de susceptibles (recordemos que  $S_{\infty} > 0$ ). Consideremos entonces la primera ecuación del sistema (2.1) e integremos a ambos lados:

$$\begin{aligned} \int_0^{\infty} S'(t) dt &= -\beta \int_0^{\infty} S(t) I(t) dt \\ S_0 - S_{\infty} &= \beta \int_0^{\infty} S(t) I(t) dt \\ S_0 - S_{\infty} &\geq \beta S_{\infty} \int_0^{\infty} I(t) dt \end{aligned}$$

Lo que nos dice que  $I(t)$  es integrable en  $[0, \infty)$  y por lo tanto  $I_{\infty} = 0$ .

Notemos que en el sistema (2.1) la variable  $R$  no aparece en ninguna de las dos primeras ecuaciones, por lo que para resolver el problema basta considerar:

$$\begin{aligned} S' &= -\beta SI \\ I' &= I(\beta S - \gamma) \end{aligned} \tag{2.2}$$

Una vez resuelto, despejamos a  $R$  mediante la ecuación  $R(t) = N - S(t) - I(t)$ . Si definimos  $\frac{\gamma}{\beta} = \rho$  (*tasa de eliminación relativa*) y utilizamos el sistema (2.2) se llega a que  $\frac{I'}{S'} = -1 + \frac{\rho}{S}$ . Reescribiendo obtenemos

$$I' = -S' + \rho \frac{S'}{S}. \tag{2.3}$$

---

<sup>1</sup>Tener presente que una epidemia es un brote de una enfermedad en una población durante un período relativamente corto de tiempo.

Observemos que hay un umbral en el punto  $S = \rho$ . Si el número inicial de susceptibles es estrictamente mayor que  $\rho$ , entonces al inicio de la epidemia el número de infecciosos aumenta; mientras que si es estrictamente menor que  $\rho$ , el número de infecciosos empieza (y se mantiene) decayendo. Definamos al *número básico de reproducción* como

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

Este parámetro es importante ya que representa el promedio de infecciones secundarias producidas cuando un individuo infeccioso es introducido en una población compuesta solamente por susceptibles. A pesar de que al final la enfermedad siempre se extingue su dinámica inicial esta gobernada por este valor. Si  $R_0 > 1$  se produce un brote epidémico, es decir, la cantidad de infecciosos aumenta al principio del proceso. Mientras que si  $R_0 < 1$  los agentes infecciosos comienzan y se mantienen decreciendo. Aquí queda de manifiesto que  $R_0$  determina el umbral de transmisión de la enfermedad. Por otro lado, como ya se ha dicho antes  $\gamma I(t)$  es el número de individuos recuperados por unidad de tiempo, a tiempo  $t$ . Si asumimos que  $I$  es constante, se sigue que en el intervalo de tiempo  $1/\gamma$ , todos los infectados  $I(t)$  se recuperan. Por lo tanto,  $1/\gamma$  es en realidad la *duración promedio de la infección*.

Si volvemos a la ecuación (2.3) e integramos en ambos lados llegaremos a la siguiente expresión con valores iniciales  $(S_0, I_0)$

$$I - I_0 = -(S - S_0) + \rho \ln \frac{S}{S_0}. \quad (2.4)$$

Notemos también que  $S_0 + I_0 = N$ . Tomemos límite en la ecuación anterior y obtendremos

$$N - S_\infty + \rho \ln \frac{S_\infty}{S_0} = 0.$$

Es fácil verificar que la expresión de arriba tiene una única raíz real  $S_\infty$  en el intervalo  $(0, N]$ . Se sigue también que

$$\rho = \frac{N - S_\infty}{\ln S_0 - \ln S_\infty}. \quad (2.5)$$

Dada la importancia que tiene  $R_0$  a la hora de que una epidemia se desarrolle, una estrategia sería estimarlo y tratar de reducirlo a un valor menor que 1 mediante técnicas de control, como por ejemplo, vacunación. Ésta tiene el efecto de mover agentes inicialmente susceptibles directamente al compartimiento de los recuperados por lo que el  $S_0$  resultaría más pequeño. Por eso, si  $S_0$  y  $S_\infty$  pueden ser estimados clínicamente, entonces utilizando la ecuación (2.5) se lograría el objetivo.

Por último, si una epidemia llegara a ocurrir nos interesaría saber qué tan grave podría llegar a ser. Para esto es interesante calcular  $I_{\max}$ . Claramente si  $R_0 < 1$  entonces  $I_{\max} = I_0$ , lo cual es razonable ya que no se da ninguna epidemia. Ahora bien si estamos en el caso contrario en el que  $R_0 > 1$  entonces recordemos las ecuaciones (2.3) y (2.4). De allí podemos deducir que  $I_{\max}$  se alcanza en  $S = \rho$ . Por lo tanto,

$$I_{\max} = N - \rho + \rho \ln \frac{\rho}{S_0}.$$

También se puede calcular de forma sencilla la cantidad de infecciosos totales que hubo a lo largo de la enfermedad,

$$I_{\text{total}} = I_0 + S_0 - S_{\infty}.$$

### 2.3.2. Tasa de contactos adecuados e incidencia.

El siguiente párrafo está basado en las definiciones de [18]. Supongamos que una enfermedad se transmite sólo por contacto directo (entre un susceptible y un infeccioso). El número de individuos contactados por un infectado por unidad de tiempo se denomina *tasa de contacto de la enfermedad*, denotada por  $U(N)$  que depende del tamaño de la población  $N$ . Supongamos que la probabilidad de infección en cada contacto es  $\beta_0$ . Definimos a la función  $\beta_0 U(N)$  como la *tasa de contactos adecuados* que describe la fuerza de la infección. Un contacto adecuado es aquel que es suficiente para la transmisión de la enfermedad, entre un susceptible y un infeccioso. Dado que la fracción de los susceptibles en la población es  $S/N$  se tiene que la tasa media de contactos adecuados que tiene un individuo infectado con agentes susceptibles es  $\beta_0 U(N) S/N$  y se denomina *tasa de infección* (aquí se asume que la población se mezcla de forma homogénea y que todos los agentes tienen la misma probabilidad de contactarse). Por último, la *incidencia* de la enfermedad es el total de nuevos infecciosos contagiados por los agentes en el compartimiento  $I$  por unidad de tiempo, a tiempo  $t$ , es decir  $\beta_0 U(N) SI/N$ .

Existen dos tipos de incidencia generalmente usadas: Si  $U(N) = kN$ , esto es, la tasa de contacto es proporcional al tamaño total de la población, entonces la incidencia es  $\beta SI$ , donde  $\beta = \beta_0 k$  es llamado coeficiente de transmisión. Este tipo de incidencia se denomina *incidencia bilineal o incidencia simple de acción de masa*. Si  $U(N) = k'$ , es decir, la tasa de contacto es constante, entonces la incidencia se convierte en  $\beta SI/N$ , donde  $\beta = \beta_0 k'$ , y se llama *incidencia estándar*.

Un estudio realizado por Anderson y May [2] observa que la incidencia estándar sería más adecuada para los seres humanos o animales que viven en grupos. Una incidencia formada por  $\beta N^{\alpha} SI/N$  fue usada para modelar 5 tipos de enfermedades infecciosas en comunidades de personas con una cantidad que variaba desde 1.000 hasta 400.000 habitantes. Sus resultados mostraron que  $\alpha$  fue estimado entre 0,03 y 0,07 [1, 2], más cerca de ser 0 que 1. Este estudio muestra que, para muchas enfermedades infecciosas, el tamaño de la población (humana) donde la afección se transmitía tiene poca influencia en la incidencia. Lo cual es consistente con el concepto de que las personas se infectan a través de sus contactos diarios y que los patrones de estos contactos son independientes del tamaño de la comunidad en un país dado. En conclusión, la incidencia estándar es más adecuada que la forma bilineal para infecciones transmitidas en grupos de seres humanos, aunque la incidencia  $\beta N^{0,05} SI/N$  fue usada y confirmada por los mismos autores en [1].

### 2.3.3. Tres parámetros importantes: $R_0$ , $\sigma$ , $R$ .

Las siguientes definiciones están basadas en [19]. Recordemos que  $R_0$  está definido como el número promedio de infecciones secundarias que ocurren cuando se introduce un individuo infeccioso en una población enteramente susceptible y se denomina *el número básico de reproducción*. Asumimos implícitamente que el forastero infeccioso permanece en la comunidad durante todo el intervalo de tiempo que pueda contagiar la enfermedad y que se mezcla en la población de la misma forma en la que lo haría un nativo del lugar. *El número de contacto*  $\sigma$ , se define como la cantidad promedio de contactos adecuados que tiene un típico agente infeccioso en la duración promedio de la infección. Finalmente, *el número de reemplazo*,  $R$  se define como la cantidad de infecciones secundarias producidas por un individuo infeccioso durante el período de infecciosidad. Notemos que el número de reemplazo cambia como función del tiempo a medida que la enfermedad evoluciona luego de la invasión inicial.

Observemos que estos tres parámetros son iguales al inicio de la epidemia, cuando toda la población es susceptible, exceptuando al invasor infeccioso. A pesar de que  $R_0$  sólo está definido al comienzo de la invasión, los otros dos indicadores,  $\sigma$  y  $R$  están definidos a todo tiempo. En la mayoría de los modelos  $\sigma$  permanece constante a medida que la enfermedad se propaga y puede usarse de forma intercambiable con  $R_0$ . Sin embargo, hay algunos modelos en los que el número de contacto se hace más pequeño que el número básico de reproducción dado que se incorporan nuevas clases de agentes re-infectados con menores niveles de infectividad posterior a la invasión de la afección en la población.

El número de reemplazo  $R$  es la cantidad de casos secundarios producidos por un típico infectado, y consecuentemente luego de que la infección ha invadido la población y ya no todos son susceptibles,  $R$  será menor a  $R_0$ . Adicionalmente, después de la invasión, como la fracción de los susceptibles es menor a uno, no todos los contactos adecuados resultan en un nuevo caso. Por lo tanto, tenemos que  $R$  será menor que  $\sigma$ . Combinando estos resultados llegamos a las siguientes desigualdades:

$$R_0 \geq \sigma \geq R$$

Con igualdad de las tres cantidades al momento de la invasión. Notar que  $R_0 = \sigma$  para casi todos los modelos, y  $\sigma > R$  luego de la invasión para todos los modelos.

## 2.4. Dinámica vital.

En este modelo se considera la dinámica vital de la población, es decir, se incluyen las muertes y nacimientos. Para simplificar el análisis y que el tamaño de la población se mantenga constante supondremos que la tasa de nacimientos es igual a la de muertes. Adicional, se supone que todos los agentes nacen susceptibles y que la tasa de muerte natural es  $\mu$  en todas las clases. Este modelo generalmente se usa para analizar enfermedades que se mantienen en el tiempo (por ejemplo, 10 o 20 años) y pueden alcanzar

el estado endémico<sup>2</sup>. Es por esto que la dinámica vital es incorporada al sistema.

El diagrama de flujo sería el siguiente:

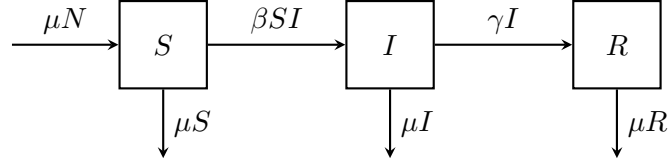


Figura 2.2: Diagrama de flujo, SIR con muertes y nacimientos.

Y su sistema de ecuaciones diferenciales asociado:

$$\begin{aligned} S' &= \mu N - \beta SI - \mu S \\ I' &= \beta SI - \gamma I - \mu I \\ R' &= \gamma I - \mu R \end{aligned} \quad (2.6)$$

Con condiciones iniciales  $S(0) = S_0$ ,  $I(0) = I_0$  y  $R(0) = 0$ .

#### 2.4.1. Propiedades matemáticas del modelo con dinámica vital.

Nuevamente dado que la variable  $R$  aparece desacoplada en el sistema (2.6) basta analizar sólo las primeras dos ecuaciones (pues  $R = N - S - I$ ).

$$\begin{aligned} \frac{dS}{dt} &= \mu N - \beta SI - \mu S := P(S, I) \\ \frac{dI}{dt} &= \beta SI - (\gamma + \mu)I := Q(S, I) \end{aligned} \quad (2.7)$$

Donde el par  $(S, I)$  pertenece a la región  $D = \{(S, I) : S \geq 0, I \geq 0, S + I \leq N\}$ . Luego dado que estamos buscando condiciones endémicas en este modelo es natural pensar qué pasa cuando el sistema se estabiliza. Esto es equivalente a pedir que todas las derivadas se anulen simultáneamente, es decir, hallar sus puntos de equilibrio. Igualando las ecuaciones de (2.6) a cero obtenemos los siguientes dos estados de equilibrio para el sistema (2.7):

$$E_0 = (N, 0) \text{ y } E_1 = \left( \frac{\mu + \gamma}{\beta}, \frac{\mu[\beta N - (\mu + \gamma)]}{\beta(\mu + \gamma)} \right) \quad (2.8)$$

A continuación nos gustaría conocer la estabilidad de cada uno de los puntos encontrados. Antes de esto notemos que la región  $D$  es positivamente invariante. Por lo tanto todas las trayectorias que comiencen dentro de esta región se mantienen en ella. Para ver esto, simple aunque tedioso, basta considerar las derivadas en los bordes. Distingamos dos casos que puedan ocurrir:

<sup>2</sup>Una endemia se refiere a un proceso patológico que se mantiene de forma estacionaria en una población o zona geográfica determinada durante períodos de tiempo prolongados.

1. Caso  $\beta N/(\mu + \gamma) < 1$  : Notemos que si queremos que  $E_1 \in D$  es necesario que  $S \leq N$ . Como  $S = (\mu + \gamma)/\beta$  debe pasar que  $(\mu + \gamma)/\beta \leq N$ , lo que es equivalente a que  $\beta N/(\mu + \gamma) \geq 1$  y en este escenario no sucede. Concluimos entonces que el único punto de equilibrio perteneciente a  $D$  es  $E_0$ . Para analizar su estabilidad planteamos el Jacobiano del sistema (2.7) en  $E_0$ :

$$J(E_0) = \begin{pmatrix} -\mu & -\beta N \\ 0 & \beta N - (\mu + \gamma) \end{pmatrix}$$

Esta matriz tiene dos autovalores:  $\lambda_1 = -\mu < 0$  y  $\lambda_2 = \beta N - (\mu + \gamma)$  que en este caso también es estrictamente menor a cero. Por ende  $E_0$  es un equilibrio asintóticamente estable. Esto nos dice que si la condición inicial está lo suficientemente cerca del punto  $E_0$ , la órbita tenderá a acercarse hacia él a medida que pase el tiempo. Sin embargo vale un resultado más fuerte y es que no es necesario que la condición inicial se encuentre cerca (basta que pertenezca a  $D$ ). Para esto notemos lo siguiente:  $D$  es positivamente invariante,  $E_0$  es su único punto de equilibrio, se encuentra en el borde y es estable. Por lo tanto no puede haber órbitas cerradas (pues si hubiera alguna, esta tendría un punto de equilibrio en su interior) ni gráficos. Por el teorema de Poincaré-Bendixson (teorema 1.2.2) se concluye que toda trayectoria con condición inicial en  $D$  tenderá asintóticamente a un punto de equilibrio (que sólo puede ser  $E_0$ ). Esto implica que sin importar la cantidad inicial de infecciosos en la población la epidemia no se puede mantener y eventualmente desaparece. El punto  $E_0$  es llamado *equilibrio libre de enfermedad*.

2. Caso  $\beta N/(\mu + \gamma) > 1$  : Notemos que ahora el segundo autovalor de  $J(E_0)$  resulta estrictamente positivo. Como consecuencia de esto  $E_0$  ya no será un equilibrio estable. Adicional, se tiene que  $E_1 \in D$ . Analicemos su estabilidad. Planteamos el Jacobiano del sistema en  $E_1$ :

$$J(E_1) = \begin{pmatrix} -\frac{\mu\beta N}{\mu+\gamma} & -(\mu + \gamma) \\ \frac{\mu[\beta N - (\mu + \gamma)]}{\mu + \gamma} & 0 \end{pmatrix}$$

Haciendo un par de cuentas se puede ver que la matriz anterior posee ambos autovalores negativos gracias a la condición  $\beta N/(\mu + \gamma) > 1$ . Se sigue que  $E_1$  es un punto de equilibrio local y asintóticamente estable.

**Proposición.** Si  $\beta N/(\mu + \gamma) > 1$  entonces  $D' = D \setminus \{(S, I) \in D : I = 0\}$  es una región asintóticamente estable para el punto de equilibrio  $E_1$  y  $\{(S, I) \in D : I = 0\}$  es una región asintóticamente estable para el punto de equilibrio  $E_0$ .

*Demostración.* Primero notar que la región  $D$  es positivamente invariante ya que ninguna trayectoria puede alejarse a través de los bordes. Toda órbita que comienza en  $D$  se mantiene allí. Por otro lado, debido a la condición  $\beta N/(\mu + \gamma) > 1$ , el punto de equilibrio  $E_0$  es un punto silla (es decir, inestable) ya que hay un



camino atractor en el eje  $S$  (autovector  $(1, 0)$  del autovalor negativo) y una dirección de repulsión hacia el interior de  $D$  con pendiente  $-1 + \gamma/\beta N$  (autovector  $(-1, 1 - \gamma/\beta N)$  de autovalor positivo). Al mismo tiempo  $E_1$  es un equilibrio local y asintóticamente estable pues posee ambos autovalores negativos. Como la región  $D$  es compacta, por el teorema de Poincaré-Bendixson (teorema 1.2.2) basta ver que no hay órbitas cerradas ni gráficos para poder concluir que cualquier trayectoria tiende asintóticamente a alguno de los dos puntos de equilibrio. Como  $E_0$  es inestable las únicas trayectorias que tenderán a él serán aquellas que comiencen en el subespacio de su autovector atractor, es decir, en  $\{(S, I) \in D : I = 0\}$ . El resto de las órbitas, aquellas en  $D'$ , tenderán a  $E_1$ .

Para ver que no hay órbitas cerradas ni gráficos aplicamos el criterio de Dulac (teorema 1.2.4) con la función  $\alpha(S, I) = 1/I$  y obtenemos

$$\frac{\partial(\alpha P)}{\partial S} + \frac{\partial(\alpha Q)}{\partial I} = -\beta - \frac{\mu}{I} < 0 \text{ para } (S, I) \in \text{int}D.$$

Una demostración más rigurosa puede hacerse mediante funciones de Liapunov. A continuación se muestra la prueba brindada por [11]. Primeramente, llamemos  $(S_e, I_e)$  al par correspondiente al estado de equilibrio  $E_1$ . Realicemos el cambio de variables  $S = S_e(1 + U)$  y  $I = I_e(1 + V)$  y obtenemos el siguiente sistema:

$$\begin{aligned} U' &= -\beta I_e U (1 + V) - \mu U - \beta I_e V \\ V' &= (1 + V) U (\mu + \gamma) \end{aligned} \tag{2.9}$$

El triángulo  $D$  se convierte entonces en

$$D^* = \{(U, V) : U \geq -1, V \geq -1, S_e U + I_e V \leq N - S_e - I_e\}.$$

El antiguo par  $(S_e, I_e)$  queda transformado con las nuevas variables en el origen,  $(0, 0)$ . Definamos la siguiente función de Liapunov  $L$  como

$$L = U^2/2 + \beta N I_e [V - \log(1 + V)] / (\mu + \gamma).$$

Esta función es positiva para  $V > -1$  y su derivada cumple

$$L' = -\beta I_e U^2(1 + V) - \mu U^2 \leq 0.$$

El conjunto  $E = \{(U, V) : L' = 0\}$  es el eje  $V$  donde  $U = 0$ . Dado que  $U = 0$  implica  $U' = -\beta I_e V$ , se puede afirmar que la única región positivamente invariante de  $E$  es el origen. En conclusión, debido al principio de Krasovskii-Lasalle (teorema 1.1.3) la sub-región de  $D^*$  que cumple  $V > -1$  es asintóticamente estable para el origen. □

La proposición anterior muestra que si  $\beta N / (\mu + \gamma) > 1$  entonces la enfermedad siempre alcanza el estado endémico a menos que la cantidad inicial de infecciosos sea nula. Esto implica que una vez que la enfermedad invade la población la epidemia persiste y su equilibrio siempre será  $E_1$ . Este punto se denomina *equilibrio endémico*.

Definamos ahora  $R_0 = \beta N / (\mu + \gamma)$ . Nuevamente se puede ver que  $R_0 = 1$  determina un umbral en el que la enfermedad se extingue ( $R_0 < 1$ ) o persiste ( $R_0 > 1$ ). Al igual que antes  $R_0$  representa la cantidad promedio de infecciones secundarias durante el curso medio de la infección producidas por la introducción de un individuo infeccioso en una población compuesta enteramente por susceptibles. Si  $R_0 < 1$ , entonces en promedio el número de nuevas infecciones por un sólo agente infeccioso a lo largo del curso medio de la infección es menor a uno, lo que implica que la enfermedad eventualmente desaparece. Por el otro lado si  $R_0 > 1$  los contagios promedios del agente infeccioso son mayores a uno, lo que lleva a la persistencia de la epidemia.

#### 2.4.2. Esperanza de vida y edad media de infección.

El siguiente párrafo está basado en las definiciones de [18]. Supongamos que  $N(a)$  es la cantidad de personas en una población que sobrevivieron hasta la edad  $a$  y que  $\mu$  es el coeficiente de la tasa de muertes naturales, es decir, la proporción de individuos que mueren en la población por unidad de tiempo. Notemos que si el tiempo y la edad tienen las mismas escalas, entonces

$$\frac{dN(a)}{da} = -\mu N(a). \quad (2.10)$$

Aquí el signo menos refleja el hecho de que  $N$  decrece con respecto a la edad debido a las muertes naturales. Asumamos que  $N(0) = N_0$ . Resolviendo la ecuación anterior se obtiene que

$$N(a) = N_0 e^{-\mu a}.$$

En efecto,  $e^{-\mu a}$  representa la probabilidad de que una persona de la población sobreviva hasta la edad  $a$ . En consecuencia, la probabilidad de muerte en el intervalo de tiempo  $[0, a]$  es  $1 - e^{-\mu a}$ . Si tomamos un individuo de forma aleatoria y denotamos a su edad de muerte como  $\xi$  entonces  $P(0 < \xi \leq a) = 1 - e^{-\mu a} = \int_0^a \mu e^{-\mu t} dt$ . Además la esperanza de la variable aleatoria  $\xi$  es

$$\mathbb{E}(\xi) = \int_0^{+\infty} t \mu e^{-\mu t} dt = \frac{1}{\mu}.$$

Por lo que  $1/\mu$  es la edad media de muerte en la población, que es lo mismo que la *esperanza de vida*.

Similarmente si  $\gamma$  es el coeficiente de la tasa de recuperación, entonces  $1/\gamma$  es la duración promedio de la infección en ausencia de muerte.  $e^{-\gamma t}$  es la probabilidad de que un individuo no se recupere hasta tiempo  $t$ . Notar sin embargo que, si se incorporan las muertes naturales, se tiene que la duración promedio de infección disminuye a  $1/(\gamma + \mu)$ .

Volvamos ahora al modelo SIR con dinámica vital (2.7) (2.8) y denotemos al equilibrio endémico como  $E_1 = (S_e, I_e)$ . Notemos que  $\beta I_e S_e$  representa el número de nuevas infecciones por unidad de tiempo cuando el sistema se encuentra en estado estable si  $R_0 > 1$ , por lo que la probabilidad de que un susceptible sea infectado por unidad de tiempo es  $\beta I_e$ . De forma similar a lo dicho anteriormente tenemos que  $e^{-\beta I_e a}$  es la

probabilidad de que los susceptibles no sean infectados hasta la edad  $a$ , entonces  $1/\beta I_e$  es la *edad media de infección* de los susceptibles o *tiempo de espera*.

Para algunas enfermedades la esperanza de vida y la edad media de infección pueden ser estimadas con datos estadísticos y usarlas para dar una estimación del número básico de reproducción. Por ejemplo en el sistema (2.6),

$$R_0 = \frac{\beta N}{\gamma + \mu} = \frac{N}{S_e} \quad \text{y} \quad S_e = \frac{N}{1 + \beta I_e / \mu}.$$

Si denotamos a la edad media de infección como  $A$  y a la esperanza de vida como  $L$  tenemos que

$$A = \frac{1}{\beta I_e} \quad \text{y} \quad L = \frac{1}{\mu}.$$

Luego, podemos reescribir a  $R_0$  como

$$R_0 = 1 + \frac{L}{A}.$$

## 2.5. Modelos con estrategias de Control.

De acuerdo con un informe de la Organización Mundial de Salud ([62], capítulo 10) la misión de minimizar la transmisión de las enfermedades infecciosas debe ser parte del núcleo de las políticas de salud pública. Las estrategias de control variarán de acuerdo a la seriedad de la enfermedad, los medios de transmisión y la tasa de contagio.

Algunas medidas para el control y la prevención de las enfermedades infecciosas son vacunación, tratamiento de la enfermedad (ej. antibióticos), profilaxis, cuarentena y aislamiento. Lo siguiente está basado en [21]:

La profilaxis es una serie de medidas para prevenir una determinada enfermedad infecciosa. Éstas pueden ser simples como higienizarse regularmente las manos con agua y jabón, o más complejas desde utilizar equipo de protección hasta tomar algún medicamento para prevenir la enfermedad.

El tratamiento es el uso de algún agente, procedimiento o régimen, como por ejemplo una droga, en un intento de curar o mitigar la afección. Hoy en día existen medicamentos para la mayoría de las infecciones que pueden curar o minimizar el impacto de la enfermedad mejorando así la calidad de vida del paciente. Por ejemplo, la malaria y la tuberculosis pueden ser tratadas con medicamentos modernos que brindan excelente tasas de recuperación. Para otras sin embargo sólo existen medicamentos que ofrecen alivio de los síntomas mas no una cura. Ejemplo de éstas incluyen al VIH y el virus del papiloma humano.

La vacunación es el proceso a través del cual se introducen microorganismos debilitados (o inactivos) dentro del cuerpo humano. Nuestro sistema inmunológico detecta a los agentes de la vacuna como externos y dispara una respuesta inmune mediante la producción de anticuerpos. Como resultado, si el mismo tipo de microorganismos

vuelven a ingresar al cuerpo serán entonces destruidos mucho más rápidos gracias a las defensas creadas. Por lo tanto un individuo que está inmunizado se encuentra protegido de la enfermedad. Si una gran mayoría de personas es vacunada, es menos probable que se produzca un brote de la enfermedad, menos aún que se propague. Este efecto es llamado *inmunidad colectiva o de rebaño*. La vacunación es uno de los grandes logros de la salud pública. Ésta ha llevado a la erradicación mundial de enfermedades como la viruela.

Sin embargo la vacunación no garantiza una protección completa ya que siempre existe la posibilidad de que la persona contraiga la enfermedad. Incluso si el huésped desarrolla anticuerpos, algunos patógenos pueden mutar (ej. virus del resfriado y de la influenza) y por ende el sistema inmune podría no ser capaz de defenderse de la infección. El grado de protección que reciben los individuos vacunados se denomina *eficacia de la vacuna*.

La cuarentena y el aislamiento son medidas de control en las que individuos expuestos a la enfermedad o infecciosos son removidos de la población para prevenir una propagación más extensa de la epidemia. La cuarentena es aplicada a individuos aparentemente sanos pero potencialmente infectados, mientras que el aislamiento se aplica a individuos ya infecciosos. El aislamiento se ha usado (y actualmente se usa) para controlar varias enfermedades peligrosas. Por lo otro lado, la cuarentena es usada menos frecuentemente pero es uno de los primeros métodos de respuesta ante una situación de extrema emergencia (por ejemplo, fue implementada durante la epidemia de SARS en 2002-2003).

El número básico de reproducción, considerado en modelos que involucran estrategias de control, depende justamente de éstas y es generalmente llamado *número básico de reproducción controlado o modificado*.

### 2.5.1. Vacunación al nacer.

Supongamos que contamos con una vacuna perfecta en donde el 100 % de las personas que se la aplica obtiene inmunidad contra la enfermedad. Asumamos que al incorporarse nueva población al sistema una fracción  $p$  de los individuos son vacunados. En consecuencia si  $\mu N$  es la tasa de reclutamiento o nacimientos, se tiene que una parte  $p\mu N$  va directo a la clase de los recuperados, y el resto  $q\mu N$  (donde  $q = 1 - p$ ) entra al compartimiento de los susceptibles. Obtenemos entonces el siguiente diagrama de flujo:

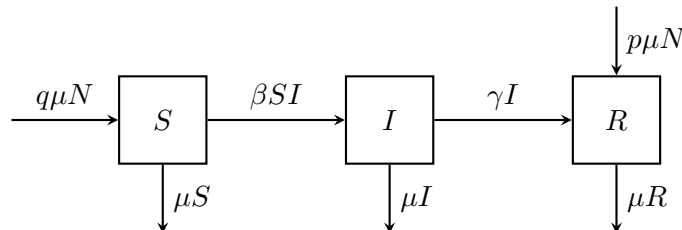


Figura 2.3: Diagrama de flujo, SIR con vacunación al nacer.

Y su sistema de ecuaciones diferenciales asociado:

$$\begin{aligned} S' &= q\mu N - \beta SI - \mu S \\ I' &= \beta SI - (\gamma + \mu)I \\ R' &= p\mu N + \gamma I - \mu R \end{aligned} \quad (2.11)$$

Con condiciones iniciales  $S(0) = S_0, I(0) = I_0$  y  $R(0) = 0$ . Al igual que en el modelo anterior el tamaño de la población total permanece constante,  $N' = 0$ . El equilibrio libre de enfermedad es obtenido igualando todas las derivadas de (2.11) a cero y asumiendo  $I = 0$ . Obtenemos que  $E_0 = (qN, 0, pN)$ . De esta forma si

$$R_0 = \frac{\beta N}{\gamma + \mu}$$

es el número básico de reproducción en la ausencia de vacunación ( $p = 0$ ), entonces  $qR_0$  es el número de reproducción en presencia de vacunación. Consecuentemente la vacunación ha reducido el número de reproducción original por la fracción  $q$ .

La pregunta ahora es ¿qué porcentaje  $p$  de la población debe vacunarse para que el número de reproducción de la enfermedad sea reducido por debajo de 1? Para esto es necesario que  $qR_0 < 1$ . Reemplazando a  $q$  por  $1 - p$  y resolviendo la desigualdad para  $p$ , se obtiene que  $p > p_c$ , donde

$$p_c = 1 - \frac{1}{R_0}.$$

Por lo tanto, si una fracción mayor a  $p_c$  es vacunada exitosamente la enfermedad no se propagará y, en efecto, toda la población estará protegida. Esto es una manifestación de la inmunidad colectiva. Recordar que ésta ocurre cuando suficientes personas han adquirido inmunidad contra la enfermedad de forma tal que la introducción de un individuo infeccioso en la población no causa una propagación de la afección.

Intuitivamente si el número de contacto es  $\sigma$  (lo que implica que una persona infectada tiene  $\sigma$  contactos adecuados durante el período infeccioso) entonces el número de remplazo  $\sigma S/N$  debe ser menor que uno para que la enfermedad no se propague. Esto significa que la fracción de susceptibles  $s = S/N$  debe ser menor que  $1/\sigma$ . Ahora bien si la fracción de inmunes debido a la vacunación es  $r = R/N$  y suponemos que  $I = 0$  entonces  $s = 1 - r$ . En consecuencia, para lograr la inmunidad colectiva por vacunación es necesario que la fracción de inmunes satisfaga

$$r > 1 - 1/\sigma = 1 - 1/R_0. \quad (2.12)$$

Por ejemplo, si  $R_0 = \sigma = 10$  entonces la fracción de inmunes debe satisfacer  $r > 1 - 1/10 = 0,9$  para que el número de remplazo sea menor a uno y la enfermedad no invada a la población.

Por varias razones una pequeña proporción de aquellos que se vacunan no desarrollan inmunidad a la enfermedad. La fracción correspondiente al fallo de la primera

aplicación de la vacuna oscila generalmente entre 0,05 y 0,10, aunque puede alcanzar los valores entre 0,2 y 0,4 para ciertas vacunas de influenza. La eficacia de la vacuna (VE) se define como la fracción de aquellos vacunados que desarrollan inmunidad. Por ejemplo, para el sarampión o la rubéola, la eficacia de las vacunas es aproximadamente 0,95. Dado que la proporción de los inmunes satisface  $r = (V)(VE)$  donde  $V$  es la fracción de la población vacunada, la desigualdad (2.12) implica que la inmunidad colectiva es lograda si la fracción de los vacunados  $V$  satisface

$$V > \frac{(1 - 1/\sigma)}{VE}.$$

La siguiente tabla extraída de [19] muestra las estimaciones estadísticas del porcentaje de personas que se deben vacunar para obtener la inmunidad colectiva para distintas enfermedades.

| Enfermedad    | Ubicación                   | $A$  | $L$ | $R_0 = \sigma = \frac{1}{1 + L/A}$ | Mín. $r$ inmun. colec. | Efect. de la vacuna | Mín. frac. a vacunar inmun. colec. |
|---------------|-----------------------------|------|-----|------------------------------------|------------------------|---------------------|------------------------------------|
| Sarampión     | Inglaterra y Gales, 1956-59 | 4.8  | 70  | 15.6                               | 0.94                   | 0.95                | 0.99                               |
|               | EE.UU., 1912-28             | 5.3  | 60  | 12.3                               | 0.92                   | 0.95                | 0.97                               |
|               | Nigeria, 1960-68            | 2.5  | 40  | 17.0                               | 0.94                   | 0.95                | 0.99                               |
| Varicella     | Maryland, EE.UU., 1943      | 6.8  | 70  | 11.3                               | 0.91                   | 0.90                | 1.01                               |
| Paperas       | Maryland, EE.UU., 1943      | 9.9  | 70  | 8.1                                | 0.88                   | 0.95                | 0.93                               |
| Rubéola       | Inglaterra y Gales, 1979    | 11.6 | 70  | 7.0                                | 0.86                   | 0.95                | 0.91                               |
|               | Alemania occidental, 1972   | 10.5 | 70  | 7.7                                | 0.87                   | 0.95                | 0.92                               |
| Poliomielitis | EE.UU., 1955                | 17.9 | 70  | 4.9                                | 0.80                   | ?                   |                                    |
|               | Países bajos, 1960          | 11.2 | 70  | 4.3                                | 0.86                   | ?                   |                                    |
| Viruela       | India                       | 12.0 | 50  | 5.2                                | 0.81                   | 0.95                | 0.85                               |

Figura 2.4: Números de reproducción y fracciones para inmunidad colectiva estimados.

### 2.5.2. Vacunación continua.

En algunos casos para ciertas enfermedades de animales salvajes o cuando se necesita impulsar la vacunación, la estrategia de control se transformará en apuntar a toda la población susceptible y no limitarse solamente a los recién nacidos [14]. En estos escenarios se modela la vacunación aleatoria de cualquier individuo, independientemente de su estado de salud, aunque sólo la vacunación de los susceptibles tiene algún efecto significativo. Introducimos un parámetro de vacunación  $v$ , que ahora representará

una tasa por unidad de tiempo en lugar de una fracción como en el caso anterior. La atención debe enfocarse entonces en la proporción de los susceptibles inmunizados por unidad de tiempo. Obtenemos así el siguiente diagrama de flujo:

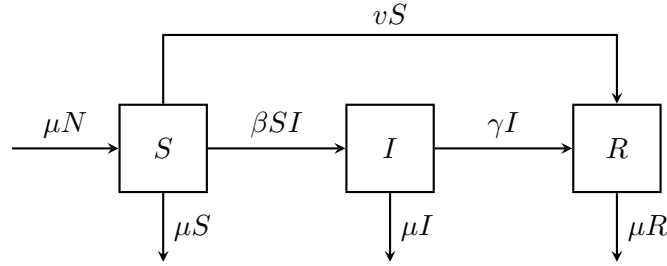


Figura 2.5: Diagrama de flujo, SIR con vacunación continua.

Y su sistema de ecuaciones diferenciales asociado:

$$\begin{aligned}
 S' &= \mu N - \beta SI - (\mu + v)S \\
 I' &= \beta SI - (\gamma + \mu)I \\
 R' &= vS + \gamma I - \mu R
 \end{aligned}
 \tag{2.13}$$

Las consecuencias dinámicas de estos cambios son cualitativamente menores. El sistema aún posee dos estados de equilibrio, uno libre de enfermedad y otro endémico. La diferencia principal es evaluar la tasa de vacunación requerida para eliminar la infección. El número básico de reproducción será ahora

$$R_c = \frac{\mu\beta N}{(\mu + v)(\mu + \gamma)} = \frac{\mu}{\mu + v} R_0.$$

Al igual que en los modelos anteriores  $R_c = 1$  define el umbral de la estabilidad de los puntos de equilibrio. La idea es entonces reducirlo a un valor menor a 1 a través de la vacunación y forzar al sistema al estado libre de enfermedad. Luego de algunas cuentas se llega al valor crítico de la tasa de vacunación,  $v_c = \mu(R_0 - 1)$ . Notemos que este criterio es claramente distinto al modelo anterior. Aquí,  $v_c$  aumenta linealmente con  $R_0$  a diferencia de la relación cóncava previa. Sin embargo, estos dos umbrales ( $p_c$  y  $v_c$ ) llevan a que sea necesario vacunar una fracción del mismo tamaño de la población para poder eliminar la infección. En el umbral crítico, se vacunan  $v_c S_e$  personas por unidad de tiempo (donde  $S_e$  es la cantidad de individuos susceptibles en el equilibrio endémico). Sustituyendo por los valores obtenemos

$$v_c S_e = \frac{\mu(R_0 - 1)N}{R_0} = \mu\left(1 - \frac{1}{R_0}\right)N = \mu p_c N.$$

En consecuencia, ambos esquemas de vacunación son equivalentes en términos del número de susceptibles que es necesario inmunizar. La diferencia clave reside en el hecho de que la vacunación continua asume que la fracción de la población susceptible

a la infección no se puede identificar sin ambigüedades y, por lo tanto, el alcance de la vacunación se extiende a toda la población a pesar de que sea efectiva solamente para el grupo susceptible. Por esta razón, controlar una enfermedad infecciosa a través de la vacunación al momento del nacimiento sería más fácil que tratar de inmunizar a toda la población susceptible.

### 2.5.3. Cuarentena y aislamiento.

Lo siguiente se basa en [12]. En este modelo se crea un compartimiento adicional  $Q$  al que pertenecen los individuos que se encuentran en cuarentena o aislados (definimos como  $Q(t)$  a la cantidad de agentes en cuarentena a tiempo  $t$ ). Para algunas enfermedades leves esto puede significar que las personas elijan quedarse en la casa y no vayan a la escuela o al trabajo porque se sienten enfermas. Para otras enfermedades más severas las personas en cuarentena pueden ser forzadas al aislamiento en hospitales o centros de control. Se asume que estos individuos no se mezclan con los demás y en consecuencia no infectan a los susceptibles. Consideremos ahora la incidencia estándar,  $\beta SI/N$ . Aquí  $\beta$  representa el promedio de contactos adecuados que tiene un agente en la población. Si suponemos que dicha persona es infecciosa entonces podría transmitir la enfermedad solamente si se contacta con un susceptible. En ausencia de cuarentena la fracción de los susceptibles sería  $S/N$ . Sin embargo en presencia de ésta como los individuos aislados no pueden ser contactados la proporción de los susceptibles aumenta a  $S/(N - Q)$ , es decir, que es más probable que el agente infeccioso contacte un susceptible. Por lo tanto, la *incidencia ajustada por cuarentena* sería  $\beta SI/(N - Q)$ .

Consideremos además que la cantidad de nacimientos no depende del tamaño de la población y es constante con tasa  $A$ . Por lo tanto  $N$  ya no resulta constante y se tiene que

$$\frac{dN}{dt} = A - \mu N \implies N \xrightarrow[t \rightarrow \infty]{} \frac{A}{\mu}.$$

Se puede incorporar también las muertes causadas por la enfermedad, donde el coeficiente de la tasa de mortandad es  $\alpha$ , es decir, que mueren  $\alpha(I + Q)$  individuos por la infección por unidad de tiempo. En este escenario  $N$  varía de la siguiente forma:

$$\frac{dN}{dt} = A - \mu N - \alpha(I + Q)$$

Asumamos adicionalmente que la media de permanencia en cuarentena sin muertes es  $1/\varepsilon$ . Consideremos que una parte de los infecciosos son puestos en aislamiento o cuarentena a tasa  $\delta I$  y que el resto permanece libre y se recuperan a tasa  $\gamma I$ . Obtenemos entonces el siguiente diagrama de flujo:



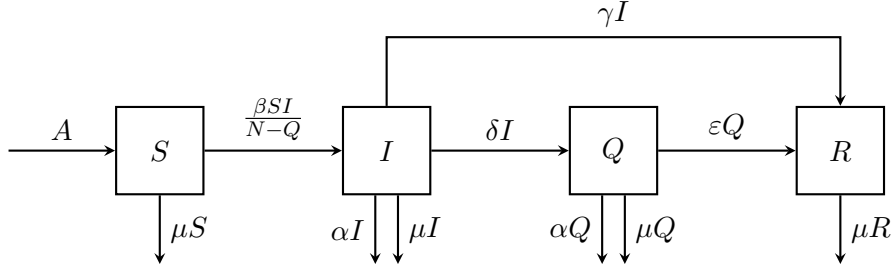


Figura 2.6: Diagrama de flujo, SIQR con incidencia ajustada y muertes por enfermedad.

Y su sistema de ecuaciones diferenciales asociado:

$$\begin{aligned}
 S' &= A - \frac{\beta SI}{N-Q} - \mu S \\
 I' &= \frac{\beta SI}{N-Q} - (\gamma + \mu + \alpha + \delta)I \\
 Q' &= \delta I - (\alpha + \mu + \varepsilon)Q \\
 R' &= \gamma I + \varepsilon Q - \mu R
 \end{aligned} \tag{2.14}$$

Es necesario observar que cuando se utiliza la incidencia estándar  $\beta$  representa la tasa de contactos adecuados por unidad de tiempo de un agente infeccioso. Si se lo multiplica entonces por el tiempo medio de permanencia en la clase  $I$  se obtendría el número básico de reproducción. En este modelo en particular el número de reproducción controlado  $R_c$  está definido de la siguiente manera:

$$R_c = \frac{\beta}{\gamma + \mu + \alpha + \delta}.$$

Notar que  $\frac{1}{\gamma + \mu + \alpha + \delta}$  es el tiempo promedio de permanencia de los individuos en el compartimiento  $I$ . Observemos que  $R_c$  disminuye a medida que el coeficiente de la tasa de cuarentena  $\delta$  aumenta. Es interesante notar que  $\varepsilon$  no aparece en  $R_c$ . Esto se debe a que el modelo asume que las personas en la clase  $Q$  no infectan a otros individuos mientras permanecen ahí y cuando se mueven afuera de  $Q$  ya no son infecciosos. Para esto último es necesario que la media de permanencia en cuarentena  $1/\varepsilon$  dure por lo menos la media del período de infección  $1/\gamma$ , es decir,  $\varepsilon \leq \gamma$ .

En [12] se analizaron seis modelos  $SIQS$  y  $SIQR$  con incidencia bilineal, estándar o ajustada a la cuarentena, y se encontró que sólo el modelo  $SIQR$  con incidencia ajustada de cuarentena podría tener una bifurcación de Hopf, en comparación con los otros cinco modelos con muerte inducida por enfermedad, cada uno de los cuales tiene un equilibrio, endémico o libre de enfermedad, globalmente estable. Además se obtuvieron condiciones necesarias y suficientes para demostrar la estabilidad de los mismos.

#### 2.5.4. Tratamiento.

Lo siguiente está basado en [18] y [41]. El tratamiento es un método importante para disminuir la propagación de enfermedades como el sarampión, la tuberculosis y

la gripe. En los modelos epidémicos clásicos se supone que la tasa de tratamiento de los enfermos es proporcional al número de infecciosos. Sin embargo esto no resulta satisfactorio debido a que los recursos para el tratamiento deberían ser bastante grandes. De hecho, cada comunidad debe tener una capacidad adecuada de tratamiento. Si es demasiado grande, la comunidad paga por un costo innecesario. Si es demasiado pequeño, la comunidad corre el riesgo del brote de una enfermedad. Por lo tanto, es importante determinar una capacidad adecuada para el tratamiento de una enfermedad.

Supongamos que la capacidad para el tratamiento de una enfermedad en una comunidad es constante y el coeficiente de tasa de recuperación debido al tratamiento es  $r$ . Supongamos además que la tasa de tratamiento es proporcional al número de los infecciosos cuando no se ha alcanzado la capacidad del tratamiento, y se satura a una constante cuando el número de infecciosos es tan grande que se excede la capacidad del tratamiento. Con estos supuestos, se construyó el siguiente modelo [41]:

$$\begin{aligned} S' &= A - \beta SI - \mu S \\ I' &= \beta SI - (\gamma + \mu)I - T(I) \\ R' &= \gamma I + T(I) - \mu R \end{aligned} \tag{2.15}$$

Donde la tasa de recuperación debido al tratamiento es:

$$T(I) = \begin{cases} rI & \text{si } 0 \leq I \leq I_0 \\ rI_0 & \text{si } I > I_0 \end{cases}$$

Notar que un tratamiento constante es adecuado cuando la proporción de los infecciosos es grande. En este modelo la tasa de tratamiento es proporcional al número de infecciosos cuando la capacidad de tratamiento no es alcanzada y, en caso contrario, toma la capacidad máxima. Esto representa, por ejemplo, la situación en la que los pacientes deben ser hospitalizados, donde el número de camas es limitado. Este escenario también es cierto para el caso en el que la cantidad de medicinas no es suficiente. La dinámica y estabilidad de este modelo, así como también el papel que juega  $I_0$ , fue estudiada en [41] por W. Wang donde se observan bifurcaciones hacia atrás en determinados umbrales.

## 2.6. Comentarios finales.

Hoy en día el mundo de los modelos epidemiológicos es muy amplio y diverso. Lo presentado en esta tesis es sólo una introducción a las herramientas más básicas de la epidemiología matemática. Sistemas y modelos más complejos han sido elaborados incorporando características que les permiten asemejarse más a la realidad como por ejemplo múltiples grupos de poblaciones, estructura de edad, migraciones y dispersión, heterogeneidad espacial y modelos con retardo, por nombrar algunos. También se han explorado los modelos discretos y estocásticos. La particularidad de estos últimos es que permiten analizar distintas estructuras, como por ejemplo redes complejas.

En este trabajo han quedado en el tintero modelos básicos en los que no existe inmunidad a la enfermedad, y por lo tanto, los agentes pueden volver a enfermarse. En ese caso, se cuenta sólo con dos compartimentos  $S, I$  y son los modelos SIS o SI. También puede considerarse un período de exposición a la enfermedad. Cuando hay un contacto adecuado entre un infectado y un susceptible de forma tal que la transmisión de la enfermedad ocurre, el agente susceptible entra en la clase de los expuestos  $E$  correspondiente a aquellos en período latente, que son infectados pero aún no infecciosos. El período de incubación está definido como aquel desde la exposición inicial hasta la aparición de los síntomas. Dado que una persona se vuelve infecciosa antes o después de que los síntomas se manifiestan, el período de incubación es a menudo distinto del de latencia. A la hora de modelar, lo importante es cuando termina la latencia ya que es en este momento en el que el agente pasa a la clase  $I$  de los infecciosos, en el sentido de que puede transmitir la infección.

La adición de un período de latencia es esencialmente similar a la introducción de un ligero *delay* en el sistema y se podría esperar que tal característica actúe para desestabilizar y ralentizar el sistema. Sin embargo las propiedades dinámicas del modelo SEIR son cualitativamente similares a las del sistema SIR. Si se considera el modelo con dinámica vital y se asume que la duración promedio del período latente viene dada por  $1/\varepsilon$  entonces la expresión para  $R_0$  ahora es ligeramente diferente, debido a la muerte de algunos individuos en la clase expuesta que no contribuyen a la cadena de transmisión. Sin embargo, esta diferencia a menudo es insignificante porque típicamente  $\varepsilon/(\mu+\varepsilon) \sim 1$  ya que el período latente es mucho más pequeño que la esperanza de vida. Como se esperaba, si el período latente es infinitesimalmente pequeño (es decir,  $\varepsilon \rightarrow \infty$ ), se recupera la misma expresión para  $R_0$  que para el modelo SIR ( $R_0 = \beta/(\gamma + \mu)$ , suponiendo incidencia estándar).

Por otro lado, podría considerarse también modelos con un compartimento  $M$  en el que se encuentran los recién nacidos que poseen inmunidad temporaria a la enfermedad gracias a los anticuerpos transmitidos por la madre a través de la placenta. Una vez que estos anticuerpos pasivos desaparecen (pues, el infante no produce nuevos) el infante se mueve a la clase de los susceptibles  $S$ .

Ejemplos de todos estos modelos son: MSEIR, MSEIRS, SEIR, SEIRS, SIR, SIRS, SEI, SEIS, SIS, SI.

El modelo SEIR será utilizado más adelante en este trabajo (capítulo 4) para modelar la epidemia de Ébola en África Occidental durante los años 2014 - 2016. En el caso del Ébola el período de latencia resulta ser igual al de incubación dado que sólo una persona que presenta síntomas puede transmitir la enfermedad.

A la hora de elegir un modelo para cierta enfermedad es necesario tener distintos factores en cuenta. Obviamente el proceso de transmisión posee varias complejidades epidemiológicas: la naturaleza del proceso de transmisión, heterogeneidades de los patógenos, patrones de contacto entre huéspedes, y períodos de incubación largos y variables de algunas enfermedades. La incorporación de estas complejidades es a menudo esencial para generar un modelo que sea cualitativamente razonable, pero también capaz de realizar predicciones cuantitativas. La mayoría de las veces se incorporan múltiples

parámetros que deben ser estimados a través de datos epidemiológicos, clínicos y de comportamiento que se encuentren disponibles.

## Capítulo 3

# Teoría de control óptimo.

Referencias principales [16], [37], [13]. Referencias secundarias [17], [3], [22], [6], [34], [35], [36].

Lo que se ha hecho hasta ahora es describir y estudiar las propiedades de distintos modelos de sistemas dinámicos. Nos hemos concentrado más que nada en la evolución a lo largo del tiempo y en hallar los umbrales que determinan las diferentes posibles evoluciones, siempre buscando que el modelo refleje ciertas propiedades cualitativas o cuantitativas de la realidad.

A continuación la idea es incorporar una función que modifique al sistema a medida que pasa el tiempo. Lo que se busca también es evaluar los beneficios o costos que resultan de dicha evolución. Supongamos entonces que  $x$  es una trayectoria o estado de un sistema dinámico. Empecemos considerando que nuestro intervalo de tiempo es  $[0, T]$  con  $T > 0$ . Supongamos además que tenemos una función de control  $u$  tal que

$$u : [0, T] \rightarrow U.$$

El conjunto  $U \subseteq \mathbb{R}$  se denominará *región de control*. Dado que  $u$  modifica al sistema es razonable que  $x$  cumpla una ecuación diferencial del estilo

$$\dot{x} = g(t, x, u) \tag{3.1}$$

donde se pone de manifiesto cierta dependencia del control. Para darle estructura al problema digamos que  $g$  es una función  $\mathcal{C}^1$  en todos sus argumentos. Admitamos solamente los controles que sean continuos a trozos y que produzcan existencia y unicidad de la soluciones de la ecuación (3.1) dotada de la condición inicial  $x(0) = x_0$ .

Una vez armado el sistema, es necesario obtener una forma de medir los beneficios o costos que produce un cierto control. Para esto incorporemos un funcional  $J$  que dependa del control  $u$ . En principio hay diversos objetos que cumplirían las propiedades de funcional pero digamos por el momento que

$$J(u) = \int_0^T f(t, x, u) dt,$$

donde  $f$  es también  $\mathcal{C}^1$  en todos sus argumentos.

Finalmente, el objetivo sería maximizar  $J$  respecto a todos los controles admisibles en el caso que  $J$  represente beneficios o bien minimizarlo cuando se habla de costos. Dado que estamos hablando de controles admisibles digamos que un control es admisible cuando pertenece a la clase  $\mathcal{A}$  donde

$$\mathcal{A} = \{u : [0, T] \rightarrow U \text{ tales que } u \text{ es continua a trozos}\}.$$

Buscamos entonces un control que sea *óptimo* en el sentido de que sea solución del problema

$$\begin{aligned} \max_{u \in \mathcal{A}} J(u) &= \int_0^T f(t, x(t), u(t)) dt \\ \text{sujeto a } \dot{x}(t) &= g(t, x(t), u(t)), \quad 0 < t < T \\ x(0) &= x_0, \quad T \text{ dado, } x(T) \text{ libre.} \end{aligned}$$

El objetivo de este capítulo es entonces estudiar las propiedades de estos problemas. La idea es abstraerse un poco de lo que representa el modelo y concentrarse sólo en los aspectos matemáticos. La literatura acerca de este tema es bastante variada y avanzada. Sin embargo, aquí sólo se presentarán las ideas fundamentales. Se podría pensar que esta teoría es avanzar un paso más en relación a la optimización no lineal en la que se buscan máximos en espacios de dimensión finita.

### 3.1. Condiciones necesarias.

Consideremos el siguiente problema de control óptimo en su forma más simple posible:

$$\begin{aligned} \max_{u \in \mathcal{A}} J(u) &= \int_0^T f(t, x(t), u(t)) dt \\ \text{sujeto a } \dot{x}(t) &= g(t, x(t), u(t)), \quad 0 < t < T \\ x(0) &= x_0, \quad T \text{ dado, } x(T) \text{ libre.} \end{aligned} \tag{3.2}$$

Asumimos que  $f$  y  $g$  son funciones continuamente diferenciables en todos sus argumentos. La clase  $\mathcal{A}$  de controles admisibles será el de las funciones continuas a trozos en el intervalo  $[0, T]$  tales que el sistema dinámico (3.2) tenga una única solución definida en  $[0, T]$ . Por el momento no hay restricciones en los valores del control por lo que la región de control es  $U = \mathbb{R}$ .

Lo que buscamos ahora son condiciones necesarias para que un control sea óptimo. Para esto supongamos primero que existe uno, sea  $u^*$  (continua a trozos) control óptimo del problema y  $x^*$  su estado dinámico asociado. En consecuencia, se tiene que  $J(u^*) \geq J(u)$  para todo control admisible. La idea ahora es introducir una perturbación y ver qué se puede deducir a medida que se la hace tender a cero. Con esto en mente, sea  $v = v(t)$  una función fija arbitraria continua a trozos en  $[0, T]$  y consideremos al conjunto de los controles de la forma

$$u(t) = u^*(t) + hv(t)$$

con  $h \in \mathbb{R}$ .  $v$  puede verse como  $\delta u$ , una variación de  $u^*$  que sería el término lineal de  $u - u^*$ . Denotemos por  $y = y(t, h)$  al estado del sistema asociado al control  $u(t) = u^*(t) + hv(t)$ . Nuevamente,  $\frac{\partial y}{\partial h} = y_h$  puede verse como  $\delta x$ , una variación de  $x$ . Observemos que  $h = 0$  implica  $u(t) = u^*(t)$  y por lo tanto

$$y(t, 0) = x^*(t), \quad y(0, h) = x_0.$$

Además, si  $|h|$  es lo suficientemente pequeño entonces  $y(t, h)$  existe en todo el intervalo  $[0, T]$ .

Si definimos

$$\phi(h) = J(u^* + hv) = \int_0^T f(t, y, u^* + hv) dt$$

entonces la condición de optimalidad se expresa como

$$\phi(0) \geq \phi(h),$$

para todo  $h$  en la vecindad del origen. Dado que  $\phi$  es una función de variable real y alcanza su máximo en  $h = 0$ , se deduce que  $\phi'(0) = 0$ . Calculemos entonces su derivada.

Sea

$$y_h(t, h) = \frac{\partial y}{\partial h}(h, t).$$

Dado que  $f$  y  $g$  son funciones regulares, se pueden derivar dentro del signo integral,

$$\phi'(h) = \int_0^T [f_x(t, y, u^* + hv) y_h(t, h) + f_u(t, y, u^* + hv) v] dt.$$

Asumamos ahora que  $h = 0$  y recordemos que  $y(t, 0) = x^*(t)$ . La expresión anterior se transforma en

$$\phi'(0) = \int_0^T [f_x(t, x^*, u^*) y_h(t, 0) + f_u(t, x^*, u^*) v] dt.$$

El mayor inconveniente ahora reside en el término  $y_h$  correspondiente al desplazamiento de la trayectoria óptima  $x^*$ , determinada por la variación  $hv(t)$ . Esta cantidad depende de  $u^*$  de forma no trivial y es difícil de calcular. Para poder removerla de la ecuación se introduce un multiplicador  $\lambda = \lambda(t)$  (de forma similar a los multiplicadores de Lagrange) y considerando la ecuación de transición  $\dot{x} - g(t, x, u) = 0$  como una restricción, se escribe el funcional como

$$J(u) = \int_0^T [f(t, x, u) + \lambda g(t, x, u) - \lambda \dot{x}] dt. \quad (3.3)$$

A continuación, se integra por partes el último término y se obtiene

$$\int_0^T \lambda \dot{x} dt = - \int_0^T \dot{\lambda} x dt + \lambda(T) x(T) - \lambda(0) x(0).$$

Volviendo a la ecuación (3.3) y reemplazando por la expresión anterior

$$J(u) = \int_0^T [f(t, x, u) + \lambda g(t, x, u) + \dot{\lambda} x] dt - \lambda(T) x(T) + \lambda(0) x(0).$$

Si en la igualdad anterior se usa que  $u(t) = u^*(t) + hv(t)$  y  $x(t) = y(t, h)$  se tiene que

$$\phi(h) = \int_0^T \left[ f(t, y, u^* + hv) + \lambda g(t, y, u^* + hv) + \dot{\lambda} y \right] dt - \lambda(T)y(T, h) + \lambda(0)x_0.$$

Derivando, esta vez se llega a

$$\phi'(h) = \int_0^T \left[ (f_x + \lambda g_x + \dot{\lambda})y_h + (f_u + \lambda g_u)v \right] dt - \lambda(T)y_h(T, h).$$

Finalmente, si se toma  $h = 0$  se sigue que

$$\phi'(0) = \int_0^T \left[ (f_x^* + \lambda g_x^* + \dot{\lambda})y_h(t, 0) + (f_u^* + \lambda g_u^*)v(t) \right] dt - \lambda(T)y_h(T, 0) = 0 \quad (3.4)$$

donde la  $*$  denota la evaluación en  $(t, x^*, u^*)$ . Observar que la nueva expresión de  $\phi'(0)$  deja de manifiesto la ventaja de haber introducido el multiplicador  $\lambda(t)$ . Es en este punto en el que se puede usar la libertad de forzar al multiplicador para hacer desaparecer a los términos  $y_h(t, 0)$  y  $y_h(T, 0)$ . Así, sea  $\lambda^*$  la solución del problema de la adjunta:

$$\begin{cases} \dot{\lambda} = -[f_x(t, x^*, u^*) + \lambda g_x(t, x^*, u^*)] & \text{en } (0, T) \\ \lambda(T) = 0. \end{cases} \quad (3.5)$$

La ecuación de la adjunta es una ecuación diferencial lineal para  $\lambda$  con coeficientes continuos. Por lo tanto existe una única solución  $\lambda^*$  definida en  $[0, T]$ . Esta elección de  $\lambda$  además cumple

$$\phi'(0) = \int_0^T (f_u^* + \lambda^* g_u^*) v dt = 0$$

para toda función continua a trozos  $v$ . En particular vale para  $v = f_u^* + \lambda^* g_u^*$ , obteniendo entonces

$$\int_0^T (f_u^* + \lambda^* g_u^*)^2 dt = 0.$$

Dado que

$$(f_u^* + \lambda^* g_u^*)^2 \geq 0$$

se deduce la ecuación de Euler-Lagrange o condición de optimalidad

$$f_u(t, x^*, u^*) + \lambda^* g(t, x^*, u^*) = 0 \quad \text{en } (0, T). \quad (3.6)$$

Estas ecuaciones forman un conjunto de condiciones necesarias que un control óptimo junto con su estado dinámico deben satisfacer. De hecho, las mismas pueden escribirse en términos del Hamiltoniano definido de la siguiente manera:

$$H(t, x, u, \lambda) = f(t, x, u) + \lambda g(t, x, u)$$

Así estas condiciones quedan establecidas en el próximo teorema.



**Teorema 3.1.1.** Sean  $u^*$  y  $x^*$  el control óptimo con su estado dinámico asociado para el problema (3.2). Existe entonces un multiplicador diferenciable a trozos  $\lambda^*$  (también conocido como variable adjunta) tal que  $u^*$ ,  $x^*$  y  $\lambda^*$  son soluciones del siguiente sistema:

$$\begin{cases} \dot{x} = g = \frac{\partial H}{\partial \lambda}, & x(0) = x_0 & (\text{dinámica}) \\ \dot{\lambda} = -(f_x + \lambda g_x) = -\frac{\partial H}{\partial x} & & (\text{ecuación de la adjunta}) \\ 0 = f_u + \lambda g_u = \frac{\partial H}{\partial u} & \text{en } u^* & (\text{condición de optimalidad}) \\ \lambda(T) = 0 & & (\text{condición de transversalidad}) \end{cases}$$

Notar que las condiciones necesarias para optimalidad establecidas en el teorema anterior son obtenidas bajo el supuesto de que el control  $u$  toma cualquier valor en  $\mathbb{R}$ . De forma análoga, las mismas condiciones son deducidas si  $u$  toma valores en un subconjunto abierto de  $\mathbb{R}$ . El caso de que el control sea acotado ( $u_{\min} \leq u \leq u_{\max}$ ) es más delicado y se verá más adelante.

El siguiente teorema, conocido como el principio del máximo, es una generalización de la condición de optimalidad y fue enunciado en 1956 por L. Pontryagin y sus estudiantes. La demostración del teorema es bastante compleja y requiere un tratamiento previo que no será brindado en esta tesis. Referimos al lector interesado al trabajo original [28] para mayor información.

**Teorema 3.1.2.** (*Principio del Máximo*). Sean  $u^*$  y  $x^*$  óptimos para (3.2), entonces existe una variable adjunta diferenciable a trozos  $\lambda(t)$  tal que

$$H(t, x^*(t), u^*(t), \lambda(t)) \geq H(t, x^*(t), u(t), \lambda(t))$$

para todos los controles admisibles  $u$  en cada tiempo  $t \in (0, T)$ . Además,  $\lambda$  cumple que

$$\begin{aligned} \dot{\lambda} &= -\frac{\partial H}{\partial x}(t, x^*, u^*, \lambda) \\ \lambda(T) &= 0. \end{aligned}$$

El teorema establece que el Hamiltoniano tiene un punto crítico, en la variable  $u$ , en  $u^*$  para cada  $t$ . No es sorprendente que este punto crítico sea máximo considerando el problema de control óptimo.

### 3.1.1. Otras consideraciones.

Hasta ahora se han analizado las condiciones necesarias para que un par  $(x^*, u^*)$  sea óptimo. Sin embargo, en la sección anterior el punto de partida asumía que existía un control óptimo. Cabe entonces preguntarse, ¿qué condiciones debe satisfacer el problema para que esto pase? Más aún, ¿qué condiciones son necesarias para que el funcional objetivo tome siempre valores finitos (incluso en el óptimo)? Por otro lado, si suponemos que tenemos un par  $(x, u)$  ¿qué condiciones son suficientes para afirmar que dicho par es óptimo? El objetivo de esta sección es contestar lo mejor posible estas preguntas aunque algunos de los teoremas más generales no serán tratados aquí ya que se perdería el rumbo de esta tesis.

También es interesante la idea de unicidad. Supongamos que existe un control óptimo  $u^*$  correspondiente a un problema de maximización. En principio,  $J(u) \leq J(u^*) < \infty$  para todos los controles admisibles en la clase  $\mathcal{A}$ . Uno diría que la unicidad de  $u^*$  estaría dada por el hecho de que si  $J(u) = J(u^*)$  entonces  $u = u^*$  salvo finitos puntos. En este caso las trayectorias asociadas serían idénticas y se denominaría *trayectoria óptima*,  $x^*$ .

Claramente, la unicidad de las soluciones del sistema de optimalidad implica la unicidad del control óptimo, si existe. La dificultad para probar la unicidad de las soluciones del sistema reside en la dirección opuesta del tiempo en las ecuaciones del estado y la adjunta. El estado tiene condiciones iniciales en el tiempo mientras que la adjunta posee condiciones finales. Sin embargo, en general, la unicidad del control óptimo no necesariamente garantiza la unicidad del sistema de optimalidad. Para demostrar la unicidad del control óptimo directamente, es necesario establecer una cierta concavidad del funcional objetivo  $J$ .

**Teorema 3.1.3.** (*Condiciones suficientes*). Supongamos que  $x^*$ ,  $u^*$  y  $\lambda^*$  satisfacen las hipótesis del teorema (3.1.1). Sea  $f$  cóncava con respecto a  $x, u$  y supongamos que vale alguna de las siguientes propiedades:

1.  $g$  cóncava en  $x, u$  y  $\lambda \geq 0$ .
2.  $g$  convexa en  $x, u$  y  $\lambda \leq 0$ .
3.  $g$  lineal en  $x, u$ .

Entonces  $x^*, u^*$  es un par óptimo para el problema (3.2).

*Demostración.* Sea  $u$  un control admisible y sea  $x$  su correspondiente trayectoria. Gracias a las condiciones necesarias sabemos que

$$J(u) - J(u^*) = \int_0^T [f(t, x, u) - f(t, x^*, u^*)] dt \leq 0. \quad (3.7)$$

Dado que  $f$  es cóncava en  $x, u$  se tiene que

$$f(t, x, u) - f(t, x^*, u^*) \leq (x - x^*)f_x^* + (u - u^*)f_u^* \quad ^1$$

Además sabemos que

$$f_x^* = -\lambda^* g_x^* - \dot{\lambda}^*, \quad f_u^* = -\lambda^* g_u^*.$$

Sustituyendo todo esto en la ecuación (3.7)

$$J(u) - J(u^*) \leq - \int_0^T \lambda^* [(x - x^*)g_x^* + (u - u^*)g_u^*] dt - \int_0^T \dot{\lambda}^* (x - x^*) dt.$$

---

<sup>1</sup>Aquí nuevamente la  $*$  denota la evaluación en  $x^*, u^*$ .

Luego integramos por partes el último término de la expresión anterior,

$$\begin{aligned}\int_0^T \dot{\lambda}^*(x - x^*) dt &= \lambda(x - x^*) \Big|_0^T - \int_0^T \lambda^*(\dot{x} - \dot{x}^*) dt. \\ &= - \int_0^T \lambda^*(g - g^*) dt.\end{aligned}$$

Finalmente,

$$J(u) - J(u^*) \leq \int_0^T \lambda^* [g - g^* - (x - x^*)g_x^* - (u - u^*)g_u^*] dt \leq 0.$$

Aquí la última desigual vale dado que la integral es no positiva<sup>2</sup> en los casos 1. y 2., mientras que es cero en el 3.  $\square$

El siguiente teorema es un ejemplo de condiciones para afirmar la existencia de un control óptimo. La prueba del mismo no es sencilla (ni corta) por lo que se insta al lector interesado a verla en [8] donde también podrá encontrar otros teoremas que garanticen existencia.

**Teorema 3.1.4.** (*Existencia*) Supongamos que la clase  $\mathcal{A}$  de los controles admisibles para el problema (3.2) incluya al conjunto de las funciones Lebesgue integrables (en lugar de solamente continuas a trozos) en  $0 \leq t \leq T$  a valores en  $\mathbb{R}$ . Supongamos que  $f(t, x, u)$  es convexa en  $u$  y que existen constantes  $C_4$  y  $C_1, C_2, C_3 > 0$  y  $b > 1$  tales que

$$\begin{aligned}g(t, x, u) &= \alpha(t, x) + \beta(t, x)u \\ |g(t, x, u)| &\leq C_1(1 + |x| + |u|) \\ |g(t, x_1, u) - g(t, x, u)| &\leq C_2|x_1 - x|(1 + |u|) \\ f(t, x, u) &\geq C_3|u|^b - C_4\end{aligned}$$

para todo  $t \in [0, T]$  y  $x, x_1, u$  en  $\mathbb{R}$ . Entonces existe un control óptimo  $u^*$  que maximiza a  $J(u)$ , con  $J(u^*) < \infty$ .

Antes de continuar notemos una sutileza del teorema anterior. Dado que ahora se admiten funciones Lebesgue integrables hay que modificar lo que se entiende por solución de la ecuación diferencial. En este caso para que  $x = x(\cdot)$  sea solución del sistema dinámico es necesario que

$$x(t) = x(0) + \int_0^t g(s, x(s), u(s)) ds$$

para todo  $t \in [0, T]$ . Si  $x$  cumple la ecuación anterior entonces es absolutamente continua en  $[0, T]$  y está unívocamente determinada por la función de control  $u$  y el dato inicial  $x(0)$ .

---

<sup>2</sup>Notar que los casos 1. y 2. son equivalentes a pedir que  $\lambda^*g$  sea cóncava en  $x, u$ .

### 3.1.2. Controles con restricciones.

En muchas aplicaciones es necesario restringir los valores que pueda tomar el control, por ejemplo,  $u \geq 0$ , o incluso  $u_{\min} \leq u \leq u_{\max}$ . En ocasiones hay ciertas limitaciones físicas que se deben respetar ya que no tendría sentido que el control tome valores fuera de dicho rango. Los ejemplos son variados en modelos físicos, sociales, médicos y/o biológicos.

Analicemos a continuación el siguiente problema:

$$\begin{aligned} \max_{u \in \mathcal{A}} J(u) &= \int_0^T f(t, x, u) dt \\ \text{sujeto a } \dot{x} &= g(t, x, u), \quad x(0) = x_0 \\ a &\leq u(t) \leq b, \quad \forall t \in [0, T] \end{aligned} \quad (3.8)$$

Aquí la clase de los controles admisibles será  $\mathcal{A} = \{u : [0, T] \rightarrow \mathbb{R} / u \text{ es continua a trozos y } a \leq u(t) \leq b \forall t \in [0, T]\}$ . Al igual que lo hecho anteriormente consideraremos una variación del control  $u = u^* + h \delta u$ , con la diferencia que ahora pediremos  $h \geq 0$ , por lo que el signo de  $u - u^*$  está determinado por el signo de  $\delta u$ . Agregando la restricción  $\dot{x} = g(t, x, u)$  a la integral junto con un multiplicador e integrando por partes al igual que antes, podemos calcular la variación  $\delta J$ , que resulta de la parte lineal de  $J - J^*$ .

$$\delta J = \int_0^T \left[ (f_x + \lambda g_x + \dot{\lambda}) \delta x + (f_u + \lambda g_u) \delta u \right] dt - \lambda(T) \delta x(T) \quad (3.9)$$

Elijamos  $\lambda$  que cumpla

$$\dot{\lambda} = -(f_x + \lambda g_x), \quad \lambda(T) = 0,$$

por lo que (3.9) se reduce a

$$\delta J = \int_0^T (f_u + \lambda g_u) \delta u dt$$

Para que  $x, u$  y  $\lambda$  sean soluciones de nuestro problema es necesario que maximicen el valor de  $J$  y por lo tanto,

$$\delta J = \int_0^T (f_u + \lambda g_u) \delta u dt \leq 0 \quad (3.10)$$

para todas las variaciones admisibles  $\delta u$ . Estas variaciones deben garantizar que  $u = u^* + h \delta u$  cumpla con las cotas de la clase  $\mathcal{A}$ . Si el control óptimo se encuentra en la cota inferior  $a$  para algún  $t$  entonces el control modificado no puede ser menor a este valor ya que resultaría inadmisibles y por lo tanto es necesario que  $\delta u \geq 0$ . Similarmente, si el control óptimo se encuentra en la cota superior  $b$ , cualquier modificación admisible debe verificar  $\delta u \leq 0$ .

En resumen:

$$\begin{aligned} \delta u &\geq 0 && \text{cuando } u^* = a, \\ \delta u &\leq 0 && \text{cuando } u^* = b, \\ \delta u &\text{ irrestricto} && \text{cuando } a < u^* < b. \end{aligned} \quad (3.11)$$

Es necesario que la desigualdad (3.10) se verifique para todos los controles modificados que cumplen con la condición anterior. Es decir,  $u^*$  será elegido de forma tal que

$$\begin{aligned} u^*(t) = a & \quad \text{sólo si } f_u + \lambda g_u \leq 0 & \quad \text{en } t, \\ a < u^*(t) < b & \quad \text{sólo si } f_u + \lambda g_u = 0 & \quad \text{en } t, \\ u^*(t) = b & \quad \text{sólo si } f_u + \lambda g_u \geq 0 & \quad \text{en } t. \end{aligned} \quad (3.12)$$

Por ejemplo, si  $u^*(t) = a$  entonces por (3.11) es necesario que  $\delta u \geq 0$ , y por lo tanto  $(f_u + \lambda g_u) \delta u \leq 0$  sólo si  $f_u + \lambda g_u \leq 0$ . Similarmente, si  $u^*(t) = b$  debe ocurrir que  $\delta u \leq 0$  para que sea una modificación admisible y por ende  $(f_u + \lambda g_u) \delta u \leq 0$  sólo si  $f_u + \lambda g_u \geq 0$ . Por último, al igual que sin restricciones, si  $a < u^*(t) < b$  entonces  $\delta u$  puede tomar cualquier signo, por lo que la condición  $(f_u + \lambda g_u) \delta u \leq 0$  sólo puede asegurarse si  $f_u + \lambda g_u = 0$ . Una afirmación equivalente a (3.12) sería:

$$\begin{aligned} f_u + \lambda g_u < 0 & \quad \Rightarrow u^*(t) = a, \\ f_u + \lambda g_u = 0 & \quad \Rightarrow a \leq u^*(t) \leq b, \\ f_u + \lambda g_u > 0 & \quad \Rightarrow u^*(t) = b. \end{aligned} \quad (3.13)$$

Es necesario notar también que las condiciones necesarias para el estado y la adjunta permanecen inalteradas,

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial \lambda}, & x(0) &= x_0, \\ \dot{\lambda} &= -\frac{\partial H}{\partial x}, & \lambda(T) &= 0. \end{aligned}$$

Otra forma de atacar el problema es a través de los multiplicadores de Karush-Kuhn-Tucker. Este procedimiento es análogo al de optimización no lineal. Se construye un Hamiltoniano modificado incorporando dos multiplicadores  $\mu_1, \mu_2$  de la siguiente forma:

$$\hat{H} = f + \lambda g + \mu_1(b - u) + \mu_2(u - a).$$

Luego, la condición de optimalidad (o maximalidad) toma la forma complementaria

$$\begin{aligned} f_u + \lambda g_u - \mu_1 + \mu_2 &= 0, \\ \mu_1 \geq 0, \quad \mu_1(b - u) &= 0, \\ \mu_2 \geq 0, \quad \mu_2(u - a) &= 0. \end{aligned}$$

Modificando adecuadamente los argumentos para los teoremas 3.1.1 y 3.1.3 es posible demostrarlos para controles con restricciones.

Esta técnica se puede generalizar a problemas con varios controles y variables de estado. Supongamos que tenemos  $u_1, \dots, u_m$  controles. Entonces consideremos  $r$  restricciones de la forma  $w_j(t, u_1, \dots, u_m) \geq 0$  donde cada una de estas funciones son por lo menos continuamente diferenciables en todos sus argumentos. Agregando los multiplicadores  $\mu_j$  el Hamiltoniano modificado sería

$$\hat{H} = f + \lambda g + \sum_{j=1}^r \mu_j w_j(t, u_1, \dots, u_m),$$

donde es necesario además que

$$\mu_j \geq 0, \quad \mu_j w_j(t, u_1, \dots, u_m) = 0 \quad \forall j = 1, \dots, r.$$

Hay casos también en lo que las restricciones del control dependen del tiempo y de las trayectorias, sin embargo, no describiremos estos casos aquí y remitimos al lector interesado para mayor información a [13] (cap. 2, sec. 11).

### 3.2. Principio de optimalidad.

Este principio fue enunciado por R. Bellman y establece que: *“Una política óptima tiene la propiedad de que, independientemente del estado inicial y la decisión inicial, las decisiones restantes deben constituir una política óptima con respecto al estado resultante de la primera decisión”*.

El principio aplica a una gran variedad de sistemas cuya evolución determinística o estocástica es influenciada sólo por el estado presente, independientemente de su historia o pasado. Además, sentó las bases de lo que hoy se denomina programación dinámica.

En nuestro contexto el mismo principio puede ser formulado de manera equivalente, dividiendo una trayectoria óptima en dos partes, una inicial y una final. *“La segunda parte de una trayectoria óptima es también una trayectoria óptima”*. Gracias a esto se obtiene el siguiente teorema:

**Teorema 3.2.1.** Sean  $u^*$  y  $x^*$  el control óptimo con su estado dinámico asociado para el problema

$$\begin{aligned} \max_{u \in \mathcal{A}} J(u) &= \int_0^T f(t, x(t), u(t)) dt \\ \text{sujeto a } \dot{x}(t) &= g(t, x(t), u(t)), \quad 0 < t < T \\ x(0) &= x_0 \quad T \text{ dado, } x(T) \text{ libre.} \end{aligned} \quad (3.14)$$

Sea  $\hat{t}$  un punto fijo en el tiempo tal que  $0 < \hat{t} < T$ . Entonces las funciones restringidas  $\hat{u}^* = u^*|_{[\hat{t}, T]}$ ,  $\hat{x}^* = x^*|_{[\hat{t}, T]}$ , forman un par óptimo para el problema restringido

$$\begin{aligned} \max_{u \in \hat{\mathcal{A}}} \hat{J}(u) &= \int_{\hat{t}}^T f(t, x(t), u(t)) dt \\ \text{sujeto a } \dot{x}(t) &= g(t, x(t), u(t)), \quad \hat{t} < t < T \\ x(\hat{t}) &= x^*(\hat{t}) \quad T \text{ dado, } x(T) \text{ libre.} \end{aligned} \quad (3.15)$$

Además, si  $u^*$  es el único control óptimo para (3.14) entonces  $\hat{u}^*$  es el único control óptimo para (3.15).

*Demostración.* Supongamos que  $\hat{u}^*$  no es óptimo, es decir, existe un control  $\hat{u}_1$  en el intervalo  $[\hat{t}, T]$  tal que  $\hat{J}(\hat{u}_1) > \hat{J}(\hat{u}^*)$ . El siguiente paso es construir un control en todo

el intervalo  $[0, T]$  de la siguiente manera:

$$u_1(t) = \begin{cases} u^*(t) & \text{si } 0 \leq t \leq \hat{t} \\ \hat{u}_1(t) & \text{si } \hat{t} < t \leq T \end{cases}$$

Sea  $x_1$  el estado asociado al control  $u_1$ . Notar que  $u_1$  y  $u^*$  coinciden en  $[0, \hat{t}]$ , por lo que  $x_1$  y  $x^*$  también deben coincidir allí. En consecuencia

$$\begin{aligned} J(u_1) - J(u^*) &= \left( \int_0^{\hat{t}} f(t, x_1, u_1) dt + \hat{J}(\hat{u}_1) \right) - \left( \int_0^{\hat{t}} f(t, x^*, u^*) dt + \hat{J}(\hat{u}^*) \right) \\ &= \hat{J}(\hat{u}_1) - \hat{J}(\hat{u}^*) > 0 \end{aligned}$$

Esto contradice el hecho de que  $u^*$  sea óptimo para (3.14). Por último, la prueba de la unicidad es análoga a lo hecho anteriormente.  $\square$

Sin embargo, es necesario notar que la generalización del principio de optimalidad a “cualquier parte de una trayectoria óptima es también una trayectoria óptima” es falsa. Por ejemplo, la parte inicial de una trayectoria óptima podría no serlo si se considera sólo dicho intervalo de tiempo. Para ilustrar esto consideremos el comportamiento de un competidor de maratón. Supongamos que debe recorrer una distancia de 40 km en el menor tiempo posible (el objetivo) y por lo tanto, debe escoger la mejor distribución de la energía (el control). Es claro que el objetivo no es cubrir una parte del trayecto en el menor tiempo posible y menos aún la parte inicial del mismo.

### 3.2.1. La ecuación de Bellman.

La aplicación del principio de optimalidad lleva a una ecuación muy interesante denominada la ecuación de Bellman. Consideremos el siguiente problema:

$$\begin{aligned} \max_{u \in \mathcal{A}} J(u) &= \int_0^T f(t, x, u) dt + \psi(x(T)) \\ \text{sujeto a } \dot{x} &= g(t, x, u), \quad 0 < t < T \\ x(0) &= \xi, \quad T \text{ dado, } x(T) \text{ libre.} \end{aligned} \tag{3.16}$$

Aquí la única modificación es haber introducido el término  $\psi(x(T))$  en el funcional objetivo. Esto no presenta grandes dificultades y mayormente el único cambio que incorpora en lo que se ha hecho hasta ahora es que la condición de transversalidad cambia a  $\lambda(T) = \psi'(x(T))$ . Asumimos que  $f, g$  y  $\psi$  son funciones continuamente diferenciables en todos sus argumentos. La clase de los controles admisibles es  $\mathcal{A} = \mathcal{A}_{[0, T]}$  dada por las funciones continuas a trozos,  $u : [0, T] \rightarrow U$ , donde  $U$  es un abierto de  $\mathbb{R}$  y tales que existe una única solución de la ecuación diferencial para  $x$ .

Sea  $V = V(t_0, x_0)$  el valor máximo de  $J$  con respecto a las condiciones iniciales

$x(t_0) = x_0$ , es decir,

$$\begin{aligned} V(t_0, x_0) &= \max_{u \in \mathcal{A}_{[t_0, T]}} \left[ \int_{t_0}^T f(t, x, u) dt + \psi(x(T)) \right] \\ &= \int_{t_0}^T f(t, x^*, u^*) dt + \psi(x^*(T)), \end{aligned}$$

donde la segunda igualdad vale gracias al teorema (3.2.1) notando el abuso de notación  $u^* = \hat{u}^*$  y  $x^* = \hat{x}^*$  y considerando la misma dinámica  $\dot{x} = g(t, x, u)$ . La función  $V$  es denominada *la función valor* y está definida para todo  $t_0 \in [0, T]$  y para cualquier estado admisible  $x_0$ . En particular,

$$V(T, x(T)) = \psi(x(T)). \quad (3.17)$$

**Teorema 3.2.2.** Si  $V$  es continuamente diferenciable con respecto a  $t$  y  $x$  entonces satisface la siguiente ecuación diferencial en derivadas parciales:

$$-\frac{\partial V}{\partial t}(t, x) = \max_{u \in U} \left[ f(t, x, u) + \frac{\partial V}{\partial x}(t, x) g(t, x, u) \right]. \quad (3.18)$$

Además,  $V$  cumple la condición final (3.17).

*Demostración.* Sea  $h > 0$  lo suficientemente pequeño. Luego partimos la integral que define a  $V$  de la siguiente manera:

$$\begin{aligned} V(t_0, x_0) &= \max_{u \in \mathcal{A}_{[t_0, T]}} \left[ \int_{t_0}^{t_0+h} f(t, x, u) dt + \int_{t_0+h}^T f(t, x, u) dt + \psi(x(T)) \right] \\ &= \int_{t_0}^{t_0+h} f(t, x^*, u^*) dt + \int_{t_0+h}^T f(t, x^*, u^*) dt + \psi(x^*(T)) \end{aligned}$$

Aplicando el principio de optimalidad (teorema 3.2.1) se deduce que el control  $u^*(t)$ , con  $t_0 + h \leq t \leq T$ , debe ser óptimo para el problema con condiciones iniciales  $x(t_0 + h) = x^*(t_0 + h)$ . Por lo tanto, se verifican las siguientes igualdades:

$$\int_{t_0+h}^T f(t, x^*, u^*) dt + \psi(x^*(T)) = V(t_0 + h, x^*(t_0 + h)) \quad (3.19)$$

$$V(t_0, x_0) = \int_{t_0}^{t_0+h} f(t, x^*, u^*) dt + V(t_0 + h, x^*(t_0 + h)) \quad (3.20)$$

En este punto notemos el hecho (no menor) de que  $x^*(t_0 + h)$  depende sólo de los valores del control  $u^*(t)$  en el intervalo  $[t_0, t_0 + h]$ , a través de la ecuación de transición ( $\dot{x} = g$ ). En consecuencia, todos los términos del lado derecho en la ecuación (3.20) dependen solamente de estos valores de  $u^*$ . Cualquier otro control  $u$ , en el mismo intervalo  $[t_0, t_0 + h]$ , determina el valor de su correspondiente trayectoria  $x(t_0 + h)$  a través de la ecuación de transición. Se puede calcular entonces

$$\int_{t_0}^{t_0+h} f(t, x, u) dt + V(t_0 + h, x(t_0 + h)) \quad (3.21)$$



cuyo valor no puede ser mayor que  $V(t_0, x_0)$ , gracias a la optimalidad de  $u^*$ . Por ende, la elección de  $u(t)$  en el intervalo  $[t_0, t_0 + h]$  debe maximizar la expresión (3.21). Esto implica que

$$V(t_0, x_0) = \max_{u \in \mathcal{A}_{[t_0, t_0+h]}} \left[ \int_{t_0}^{t_0+h} f(t, x, u) dt + V(t_0 + h, x(t_0 + h)) \right], \quad (3.22)$$

donde el máximo se busca con  $u$  perteneciente a la clase  $\mathcal{A}_{[t_0, t_0+h]}$  donde los controles admisibles están definidos en  $[t_0, t_0 + h]$ , es decir,  $u : [t_0, t_0 + h] \rightarrow U$ .

Dado que  $V$  es diferenciable, se tiene la siguiente expresión

$$V(t_0 + h, x(t_0 + h)) = V(t_0, x_0) + \frac{\partial V}{\partial t}(t_0, x_0) h + \frac{\partial V}{\partial x}(t_0, x_0)(x(t_0 + h) - x_0) + o(h).$$

Juntando esto con la ecuación (3.22), se deduce

$$0 = \max_{u \in \mathcal{A}_{[t_0, t_0+h]}} \left[ \int_{t_0}^{t_0+h} f(t, x, u) dt + \frac{\partial V}{\partial t}(t_0, x_0) h + \frac{\partial V}{\partial x}(t_0, x_0)(x(t_0 + h) - x_0) + o(h) \right].$$

Haciendo ahora tender  $h$  a cero, la elección de  $u$  en  $\mathcal{A}_{[t_0, t_0+h]}$  se reduce a elegir  $u$  en  $t_0$ , es decir, elegir un punto de  $U$ . En consecuencia, dividiendo por  $h$  y pasando al límite<sup>3</sup>  $h \rightarrow 0$ , se tiene que

$$0 = \max_{u \in U} \left[ f(t_0, x_0, u) + \frac{\partial V}{\partial t}(t_0, x_0) + \frac{\partial V}{\partial x}(t_0, x_0) \dot{x}(t_0) \right].$$

Lo que se hace ahora es buscar el máximo de una función real en un conjunto  $U \subseteq \mathbb{R}$ . Finalmente, utilizando nuevamente la ecuación de transición, se obtiene

$$-\frac{\partial V}{\partial t}(t_0, x_0) = \max_{u \in U} \left[ f(t_0, x_0, u) + \frac{\partial V}{\partial x}(t_0, x_0) g(t_0, x_0, u) \right]$$

que es la ecuación (3.18) evaluada en  $(t_0, x_0)$ . □

La ecuación (3.18) se conoce como la ecuación de Bellman - Hamilton - Jacobi y es uno de los principales resultados de la programación dinámica. Notemos que la expresión que se maximiza en el lado derecho de la ecuación es similar al Hamiltoniano pero con  $\frac{\partial V}{\partial x}$  en el lugar de  $\lambda$ . Resolviendo este problema de maximización se obtiene una función  $w$  que depende de  $t$  y  $x$ . Además, a partir de  $w$  es posible construir el control óptimo y su correspondiente trayectoria.

**Teorema 3.2.3.** Supongamos que estamos en las hipótesis del teorema 3.2.2,  $V$  es la solución de la ecuación de Bellman y cumple con la condición final (3.17). Adicionalmente, sea

$$w = w(t, x) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$$

una función continua a trozos respecto de  $t$  y continuamente diferenciable respecto a  $x$  tal que cumple la siguiente ecuación:

$$f(t, x, w(t, x)) + \frac{\partial V}{\partial x}(t, x) g(t, x, w(t, x)) = \max_{u \in U} \left[ f(t, x, u) + \frac{\partial V}{\partial x}(t, x) g(t, x, u) \right].$$

---

<sup>3</sup>En estas condiciones, el límite del máximo es el máximo del límite.

Básicamente, para cada  $(t, x)$  la función  $w$  toma el valor  $u \in U$  que maximiza el término derecho de la ecuación de Bellman. En consecuencia, la solución de

$$\begin{cases} \dot{x} = g(t, x, w(t, x)), & \text{con } 0 < t < T \\ x(0) = \xi \end{cases} \quad (3.23)$$

es la trayectoria óptima del problema original y

$$u(t) = w(t, x(t))$$

es el control óptimo.

*Demostración.* Sea  $u$  un control admisible y  $x$  su estado asociado. Por otro lado, sea  $x^*$  la solución de (3.23) y sea  $u^*(t) = w(t, x^*(t))$ . Para probar el teorema hay que ver que  $J(u^*) \geq J(u)$ . Gracias a la ecuación de Bellman se tiene que

$$\begin{aligned} V(T, x^*(T)) - V(0, \xi) &= \int_0^T \frac{d}{dt} V(t, x^*(t)) dt \\ &= \int_0^T \left[ \frac{\partial V}{\partial t}(t, x^*) + \frac{\partial V}{\partial x}(t, x^*) \frac{dx^*}{dt} \right] dt \\ &= - \int_0^T f(t, x^*, u^*) dt. \end{aligned} \quad (3.24)$$

De forma análoga,

$$\begin{aligned} V(T, x(T)) - V(0, \xi) &= \int_0^T \left[ \frac{\partial V}{\partial t}(t, x) + \frac{\partial V}{\partial x}(t, x) \frac{dx}{dt} \right] dt \\ &\leq - \int_0^T f(t, x, u) dt. \end{aligned} \quad (3.25)$$

Ahora, restando a (3.24) la desigualdad (3.25) y utilizando la condición final (3.17) que verifica  $V$  se deduce que

$$\psi(x^*(T)) - \psi(x(T)) \geq \int_0^T f(t, x, u) dt - \int_0^T f(t, x^*, u^*) dt,$$

que resulta equivalente a

$$J(u^*) \geq J(u).$$

□

### 3.2.2. Interpretación de la adjunta.

Existe una relación entre la variable adjunta  $\lambda$  y la derivada con respecto al estado de la función valor dada por

$$\frac{\partial V}{\partial x}(t_0, x_0) = \lim_{h \rightarrow 0} \frac{V(t_0, x_0 + h) - V(t_0, x_0)}{h} = \lambda^*(t_0).$$

En el caso de que el funcional objetivo represente un costo o ganancia se tiene que la magnitud de  $\frac{\partial V}{\partial x}$  es dinero por unidad de ítem. En consecuencia, la variable adjunta  $\lambda^*(t_0)$  es igual a la variación marginal en la función valor con respecto al estado a tiempo  $t_0$ . Se podría decir que  $\lambda^*$  es el dinero adicional (costo/ganancia) asociado a un incremento adicional en la variable estado, en un plan óptimo. En particular esta interpretación es válida para todo tiempo  $t$  (siempre que  $\frac{\partial V}{\partial x}$  exista, detalle no trivial),

$$\frac{\partial V}{\partial x}(t, x^*(t)) = \lambda^*(t).$$

Se podría hacer la siguiente aproximación,

$$V(t_0, x_0 + h) \approx V(t_0, x_0) + h \lambda^*(t_0)$$

suponiendo que  $h > 0$  es pequeño. Sin embargo, sólo con el propósito de continuar con la idea tomemos  $h = 1$ , lo significaría que

$$V(t_0, x_0 + 1) \approx V(t_0, x_0) + \lambda^*(t_0).$$

Esto podría interpretarse como que la ganancia o costo adicional que resulta de agregar una unidad al estado a tiempo inicial viene dado por  $\lambda^*(t_0)$ . Por ejemplo en una fábrica representaría la ganancia adicional asociada al incremento en una unidad de materia prima. En ciertas aplicaciones esta interpretación ayudaría a conocer qué signo (positivo o negativo) se puede esperar a partir de la variable adjunta.

### 3.3. El Hamiltoniano.

Anteriormente hemos visto que los problemas de control óptimo pueden ser enunciados, gracias al principio del máximo, de manera equivalente utilizando al Hamiltoniano. En principio  $H$  es una función de cuatro variables  $(t, x, u, \lambda)$ . Si se considera que  $x, u, \lambda$  también dependen de  $t$  se podría pensar que  $H$  es implícitamente una función de  $t$ . Por otro lado,  $x$  y  $\lambda$  son continuas y  $u$  es continua a trozos, por lo que esto basta para afirmar que  $H$  es continuo a trozos. Sin embargo, el Hamiltoniano posee dos propiedades más fuertes que serán descritas en los siguientes teoremas.

**Definición 3.3.1.** Si  $f$  y  $g$  no dependen explícitamente de  $t$ , es decir,  $f = f(x, u)$ ,  $g = g(x, u)$  diremos que es un *problema autónomo*.

**Teorema 3.3.1.** El Hamiltoniano es una función Lipschitz en la variable  $t$  sobre la trayectoria óptima.

*Demostración.* Sea  $u^*, x^*$  un par óptimo para el problema de (3.2) y sea  $\lambda^*$  la variable adjunta asociada. Sea  $t \in [0, T]$  y definamos a

$$M(t) = H(t, x^*(t), u^*(t), \lambda^*(t)).$$

Dado que  $u^*$  es una función continua a trozos en un intervalo compacto existe  $P$  intervalo acotado tal que  $u^* \in P$  para todo  $t \in [0, T]$ . Análogamente existen  $Q$  y  $R$  intervalos acotados tales que  $x^* \in Q$  y  $\lambda^* \in R$  para todo  $t \in [0, T]$ .

Consideremos ahora al Hamiltoniano como una función de cuatro variables independientes  $H = H(t, x, u, \lambda)$  donde pensamos que  $x, u$  y  $\lambda$  representan solamente números. Gracias a las propiedades originales de  $f$  y  $g$  se puede afirmar que  $H$  es continuamente diferenciable en todos sus argumentos. Por lo tanto existe una constante  $K_1$  que cumple simultáneamente

$$\left| \frac{\partial H}{\partial t}(t, x, u, \lambda) \right|, \left| \frac{\partial H}{\partial x}(t, x, u, \lambda) \right|, \left| \frac{\partial H}{\partial \lambda}(t, x, u, \lambda) \right| \leq K_1$$

para todas las tuplas  $(t, x, u, \lambda)$  en el conjunto acotado  $[0, T] \times Q \times P \times R$ . Sean  $s, t \in [0, T]$  fijos. Por comodidad escribiremos  $x_s = x^*(s)$  y  $x_t = x^*(t)$ . Análogamente con  $u_s, u_t, \lambda_t, \lambda_s$ . Sea  $\tau \in P$ . Aplicando el teorema del valor medio se tiene que

$$\begin{aligned} |H(t, x_t, \tau, \lambda_t) - H(s, x_s, \tau, \lambda_s)| &\leq \left| \frac{\partial H}{\partial t}(c_1, x_t, \tau, \lambda_t) \right| |t - s| + \left| \frac{\partial H}{\partial x}(s, c_2, \tau, \lambda_t) \right| |x_t - x_s| \\ &\quad + \left| \frac{\partial H}{\partial \lambda}(s, x_s, \tau, c_3) \right| |\lambda_t - \lambda_s| \\ &\leq K_1 |t - s| + K_1 |x_t - x_s| + K_1 |\lambda_t - \lambda_s|, \end{aligned}$$

donde  $c_1, c_2, c_3$  son puntos intermedios,  $c_1 \in [0, T]$ ,  $c_2 \in P$ ,  $c_3 \in Q$ . Por otro lado,  $x^*$  y  $\lambda$  son funciones diferenciables a trozos en un intervalo compacto por lo que son Lipschitz continuas. Sea  $K_2$  la máxima de las dos constantes de Lipschitz. En consecuencia,

$$\begin{aligned} |H(t, x_t, \tau, \lambda_t) - H(s, x_s, \tau, \lambda_s)| &\leq K_1 |t - s| + K_1 |x_t - x_s| + K_1 |\lambda_t - \lambda_s| \\ &\leq (K_1 + 2K_1K_2) |t - s|. \end{aligned} \tag{3.26}$$

Definamos  $K = K_1 + 2K_1K_2$  y notemos que la desigualdad anterior vale para cualquier  $\tau \in P$ .

Recordemos que  $M(t) = H(t, x_t, u_t, \lambda_t)$  y análogamente para  $s$ . Gracias al Principio del máximo (teorema 3.1.2) el Hamiltoniano se maximiza puntualmente por  $u^*$  y por lo tanto vale que

$$\begin{aligned} H(t, x_t, u_s, \lambda_t) &\leq H(t, x_t, u_t, \lambda_t), \\ H(s, x_s, u_t, \lambda_s) &\leq H(s, x_s, u_s, \lambda_s). \end{aligned} \tag{3.27}$$

Usando la expresión (3.26) para  $\tau = u_s$  y  $\tau = u_t$ , y combinándolo con (3.27), se obtiene que

$$\begin{aligned} -K|t - s| &\leq H(t, x_t, u_s, \lambda_t) - H(s, x_s, u_s, \lambda_s) \\ &\leq H(t, x_t, u_t, \lambda_t) - H(s, x_s, u_s, \lambda_s) \\ &= M(t) - M(s) \\ &\leq H(t, x_t, u_t, \lambda_t) - H(s, x_s, u_t, \lambda_s) \\ &\leq K|t - s|. \end{aligned}$$

Es decir, hemos probado que  $|M(t) - M(s)| \leq K|t - s|$  y como  $t$  y  $s$  eran arbitrarios se concluye que  $M$  es Lipschitz continuo.  $\square$

**Teorema 3.3.2.** Si un problema de control óptimo es autónomo, entonces el Hamiltoniano es una función constante del tiempo a lo largo de la trayectoria óptima.

*Demostración.* Sea  $u^*, x^*$  un par óptimo para el problema de (3.2) y sea  $\lambda^*$  la variable adjunta asociada. Sea  $M(t) = H(x^*(t), u^*(t), \lambda^*(t))$ . Notar que el Hamiltoniano depende sólo de tres variables ya que el problema es autónomo. Dado que  $M$  es Lipschitz continua se puede afirmar que es diferenciable en casi todo punto, con respecto a la medida de Lebesgue. Sea  $t_0 \in (0, T)$  un punto donde  $\frac{dM}{dt}$  existe.

Definamos  $u^*(t_0) = \tau$  y sea  $\delta > 0$  lo suficientemente chico para que  $t_0 \pm \delta \in (0, T)$ . Gracias al teorema (3.1.2) tenemos que  $M(t_0 + \delta) \geq H(x^*(t_0 + \delta), u^*(t_0), \lambda^*(t_0 + \delta)) = H(x^*(t_0 + \delta), \tau, \lambda^*(t_0 + \delta))$ . En consecuencia,

$$M(t_0 + \delta) - M(t_0) \geq H(x^*(t_0 + \delta), \tau, \lambda^*(t_0 + \delta)) - H(x^*(t_0), \tau, \lambda^*(t_0)).$$

Ahora dividamos por  $\delta$  y hagamos tender  $\delta$  a cero, así obtenemos

$$\begin{aligned} \frac{dM}{dt}(t_0) &\geq \left. \frac{d}{dt} H(x^*(t), \tau, \lambda^*(t)) \right|_{t=t_0} \\ &= \frac{\partial H}{\partial x}(x^*(t_0), \tau, \lambda^*(t_0)) \dot{x}^*(t_0) + \frac{\partial H}{\partial \lambda}(x^*(t_0), \tau, \lambda^*(t_0)) \dot{\lambda}^*(t_0) \\ &= -\dot{\lambda}^*(t_0) \dot{x}^*(t_0) + \dot{x}^*(t_0) \dot{\lambda}^*(t_0) = 0 \end{aligned}$$

Análogamente se puede ver que

$$M(t_0) - M(t_0 - \delta) \leq H(x^*(t_0), \tau, \lambda^*(t_0)) - H(x^*(t_0 - \delta), \tau, \lambda^*(t_0 - \delta)).$$

Dividiendo nuevamente por  $\delta$  y haciéndolo tender a cero, obtenemos que  $\frac{dM}{dt}(t_0) \leq 0$ . En conclusión,  $\frac{dM}{dt} = 0$  en casi todo punto y sumado al hecho de que  $M$  es continua, se prueba que  $M$  es constante.  $\square$

### 3.4. Métodos numéricos

Consideremos el problema de control óptimo:

$$\begin{aligned} \max_{u \in \mathcal{A}} J(u) &= \int_{t_0}^{t_1} f(t, x(t), u(t)) dt \\ \text{sujeto a } \dot{x}(t) &= g(t, x(t), u(t)) \\ x(t_0) &= a \end{aligned} \tag{3.28}$$

En esta sección se presentarán dos métodos para resolver el problema de forma numérica. Se desea encontrar un algoritmo que permita generar una aproximación del control óptimo (continuo a trozos)  $u^*$ . Para esto, es necesario partir el intervalo de tiempo  $[t_0, t_1]$  con los puntos intermedios  $t_0 = b_1 < b_2 < \dots < b_N < b_{N+1} = t_1$ . A menos que se diga lo contrario, asumiremos que los puntos se encuentran equiespaciados, es decir,  $b_{j+1} - b_j = (t_1 - t_0)/N = h$ . De esta forma obtendremos una aproximación a la solución que será un vector de la forma  $\vec{u} = (u_1, u_2, \dots, u_{N+1})$ , donde  $u_j \approx u(b_j)$ .

Un primer acercamiento, relativamente directo, sería el siguiente:

1. Elegir un ODE solver y una regla de cuadratura compuesta  $Q$  que utilicen los  $N + 1$  puntos  $b_j$  del intervalo  $[t_0, t_1]$ .
2. Definir la función

$$\begin{aligned} \hat{J} : \mathbb{R}^{N+1} &\rightarrow \mathbb{R} \\ \vec{u} &\mapsto Q(f) \end{aligned}$$

de la siguiente forma: Se recibe el input  $\vec{u}$ . Luego se aplica el ODE solver en la ecuación diferencial de (3.28) utilizando a  $\vec{u}$  como dato y se consigue  $\vec{x} \in \mathbb{R}^{N+1}$  que sería una aproximación de  $x$  (es decir,  $x_j \approx x(b_j)$ ). Posteriormente se aplica la regla de cuadratura compuesta  $Q$  (usando a  $\vec{u}$  y  $\vec{x}$ ) a la función  $f$ . El resultado de esta operación es el output.

3. Definir el problema de optimización no lineal

$$\min_{\vec{u}} - \hat{J}(\vec{u}). \quad (3.29)$$

4. Aplicar finalmente las técnicas de la teoría de optimización no lineal para funciones vectoriales y obtener  $\vec{u}^*$ .

Antes de continuar observemos algunas ventajas y desventajas.

- El método no requiere un trabajo algebraico previo del problema y la aplicación del mismo es relativamente directa. Esto hace que sea una buena opción para problemas en los que existen múltiples variables en el control y/o el estado, como también en el caso de que hayan restricciones sobre estos.
- Notar que si el control debe cumplir restricciones éstas deben ser incorporadas en el problema de optimización no lineal.
- Es necesario tener cuidado a la hora de afirmar la convergencia a  $u^*$ , en el sentido de que  $\vec{u}^*$  podría ser un extremo relativo mas no absoluto.
- Muchos de los métodos de optimización no lineal requieren cierta estructura de convexidad de la función objetivo. En este caso demostrar la convexidad de  $\hat{J}$  resulta no trivial.
- Evaluar a  $\hat{J}$  puede ser costoso en términos de tiempo a medida que crece la cantidad de puntos en los que se divide el intervalo. El motivo es que cada vez que se evalúa a  $\hat{J}$  es necesario resolver una ecuación diferencial y una regla de cuadratura compuesta.
- Este método es aplicable a un abanico más amplio de problemas relacionados con el cálculo variacional. Sin embargo desaprovecha la estructura específica de la teoría de control óptimo.

### 3.4.1. Método de barrido hacia delante y hacia atrás.

Otro acercamiento, en el que se resuelve de manera indirecta la cuestión, es considerar la información específica de este tipo de problemas (capítulos 4 y 8 de [16]). Como se ha dicho antes, cualquier solución de (3.28) debe satisfacer:

$$\begin{aligned} \dot{x} &= g(t, x, u), \quad x(t_0) = a \\ \dot{\lambda} &= -\frac{\partial H}{\partial x} = -(f_x(t, x, u) + \lambda g_x(t, x, u)), \quad \lambda(t_1) = 0 \\ 0 &= \frac{\partial H}{\partial u} = f_u(t, x, u) + \lambda g_u(t, x, u) \quad \text{en } u^* \end{aligned} \quad (3.30)$$

Ahora bien, la tercera ecuación (condición de optimalidad) generalmente puede manipularse para conseguir una expresión de  $u^*$  en términos de  $t, x$  y  $\lambda$ . Si usamos esta representación en las ecuaciones diferenciales de  $x$  y  $\lambda$ , entonces obtenemos un sistema de dos ecuaciones diferenciales con datos en los bordes. Existen múltiples métodos numéricos para resolver este tipo de problemas, como por ejemplo Runge-Kutta o esquemas adaptativos (para valores iniciales) o métodos de disparo (para datos en los bordes).

Observemos que para el estado  $x$  tenemos una condición inicial (a tiempo cero) mientras que para la adjunta  $\lambda$  el dato se encuentra a tiempo final. Por otro lado, al ser  $g$  una función de  $t, x$  y  $u$ , los valores de  $\lambda$  no son necesarios para resolver la ecuación diferencial de  $x$ , usando por ejemplo un ODE solver estándar.

A continuación se muestra un resumen del algoritmo del método de barrido hacia delante y hacia atrás. En lo que sigue,  $\vec{x} = (x_1, \dots, x_{N+1})$  y  $\vec{\lambda} = (\lambda_1, \dots, \lambda_{N+1})$  son aproximaciones del estado y la adjunta respectivamente.

**Paso 1:** Proponer una estimación inicial factible para  $\vec{u}$ .

**Paso 2:** Utilizando la condición inicial  $x_1 = x(t_0) = a$  y los valores de  $\vec{u}$ , obtener  $\vec{x}$  resolviendo hacia delante en el tiempo la ecuación diferencial de  $\dot{x}$  en el sistema de optimalidad (3.30).

**Paso 3:** Utilizando la condición de transversalidad  $\lambda_{N+1} = \lambda(t_1) = 0$  y los valores de  $\vec{u}$  y  $\vec{x}$ , obtener  $\vec{\lambda}$  resolviendo hacia atrás en el tiempo la ecuación diferencial de  $\dot{\lambda}$  en el sistema de optimalidad (3.30).

**Paso 4:** Actualizar el valor de  $\vec{u}$  utilizando los nuevos valores de  $\vec{x}$  y  $\vec{\lambda}$  en la caracterización del control óptimo, tercera ecuación de (3.30).

**Paso 5:** Verificar convergencia. Si se comparan los valores de las variables en esta iteración con la anterior y resultan “relativamente cercanos”, devolver como output los valores actuales de las soluciones. En caso contrario, volver al paso 2.

A continuación algunas observaciones del algoritmo:

- El método requiere un trabajo algebraico previo. A medida que el problema crece en complejidad armar el sistema de optimalidad se hace cada vez más difícil. Cuantas más variables tienen el control y el estado, más ecuaciones es necesario plantear.
- Generalmente  $\vec{u} \equiv 0$  es un buen candidato inicial. Sin embargo, hay que tener cuidado en caso que fuera necesario dividir por  $u$ . En estos escenarios debe hacerse una estimación inicial diferente.
- A veces, es necesario modificar la estimación inicial si el algoritmo tiene problemas con la convergencia.
- Notar que si el control debe cumplir restricciones, la propuesta inicial debe hacerlo también para que resulte factible. Además, es necesario incorporar estas restricciones en el paso 4. Debe alterarse la caracterización del control óptimo en la tercera ecuación del sistema (3.30) para reflejar las restricciones que debe cumplir.
- A menudo en el paso 4 es recomendable utilizar una combinación convexa de los valores de los controles anteriores con la caracterización actual. Esto podría ayudar a acelerar la convergencia.
- Para los pasos 2 y 3 se puede emplear cualquier solver estándar para ecuaciones diferenciales ordinarias. Los métodos de Runge-Kutta son buenas opciones, entre ellos Euler modificado o Heun, por nombrar algunos. En Matlab se encuentran implementados los solvers ODE23 y ODE45 que están basados en métodos de Runge-Kutta y resultan excelentes opciones a la hora de resolver estos tipos de problemas.
- Existen varios tests de convergencia para el paso 5. Uno bastante razonable es pedir que el error relativo sea pequeño en alguna norma. Por ejemplo, sería razonable pedir que

$$\frac{\|\vec{u} - ol\vec{d}u\|_1}{\|\vec{u}\|_1} \leq \delta. \quad (3.31)$$

Aquí  $\vec{u}$  es el vector de valores estimados del control en la iteración actual y  $ol\vec{d}u$  es el de la iteración anterior.  $\|\cdot\|_1$  se refiere a la norma  $\ell_1$  de vectores (es decir,  $\|\vec{u}\|_1 = \sum_{j=1}^{N+1} |u_j|$ ) y  $\delta$  representa la tolerancia aceptada. Sin embargo, un problema del test anterior es que no acepta controles nulos, por lo que hay que modificarlo ligeramente. Se multiplica a ambos lados en (3.31) por  $\|\vec{u}\|_1$  para remover el denominador. El test sería entonces

$$\delta \|\vec{u}\|_1 - \|\vec{u} - ol\vec{d}u\|_1 \geq 0.$$

### 3.5. Comentarios finales.

Como se ha dicho al inicio del capítulo sólo se han tratados los aspectos fundamentales de la teoría de control óptimo y mucho ha quedado en el tintero. Los problemas



a considerar podrían ser bastante más complejos, por ejemplo, incorporando ecuaciones en derivadas parciales para el estado dinámico o considerando restricciones para el control que dependan de los valores de las trayectorias. Por otro lado, una extensión a varias variables y controles, es relativamente natural y no presenta mayores complejidades. Los resultados descritos en este capítulo también sirven para este escenario. Una línea interesante que no se ha mencionado es considerar problemas discretos y/o estocásticos, ya que muchos fenómenos de la realidad poseen esta estructura.

Por otro lado, desde el inicio hemos considerado solamente problemas en los que todas las funciones viven en el intervalo  $[0, T]$  con  $T$  fijo de antemano. Ciertas aplicaciones requieren un horizonte  $T$  desconocido o como parámetro para optimizar. En estos casos, el Hamiltoniano se modifica incorporando un nuevo multiplicador  $\lambda_0$  (donde  $\lambda_0 = 1$  o  $\lambda_0 = 0$ ) de la siguiente manera:

$$H(t, x, u, \lambda_0, \lambda) = \lambda_0 f(t, x, u) + \lambda g(t, x, u).$$

Finalmente, es interesante notar que la teoría puede desarrollarse considerando funciones medibles Lebesgue donde la solución a una ecuación diferencial se toma en el sentido débil.



## Capítulo 4

# Ébola.

Me gustaría empezar este último capítulo con el siguiente extracto: *“La palabra Ébola evoca imágenes horribles de enfermedad y muerte. Las películas y los programas de televisión han dramatizado los resultados mortales de la infección con el virus. De hecho, el virus del Ébola se ha convertido en sinónimo de una muerte espantosa. Los órganos del cuerpo se “licúan”, los pacientes sucumben a una neblina de fiebre y dolor, el sangrado interno y externo provoca la exudación de sangre, y los pacientes a menudo mueren. Incluso las personas que no están familiarizadas con las enfermedades infecciosas probablemente hayan escuchado sobre el virus del Ébola. Al igual que la peste negra o el SIDA, el Ébola es una enfermedad que ha trascendido la medicina para convertirse en parte de la cultura popular. Y también como el SIDA, el virus del Ébola ha capturado el respeto del público en general y ha infundido miedo en un período de tiempo notablemente corto: ambas enfermedades se descubrieron hace apenas 30 años, y ya han dejado su huella en los libros de historia.”* [39]

Ebola, deadly diseases and epidemics, Tara C. Smith.

### 4.1. Características de la enfermedad.

La enfermedad por el virus del Ébola (EVE), antes llamada fiebre hemorrágica del Ébola, es una de las enfermedades virales más mortales en el ser humano. Fue descubierta en 1976 debido a dos brotes simultáneos ocurridos en diferentes partes de África Central. El primero de ellos sucedió en Yambuku una aldea de la actual República Democrática del Congo, R.D.C., (anteriormente Zaire) situada cerca del río Ébola, de donde el virus toma su nombre. El segundo brote ocurrió en lo que ahora es Sudán del Sur, aproximadamente a 850 km de distancia.

Inicialmente, se pensó que estos brotes eran un evento único debidos a una persona viajando entre los dos lugares. Posteriormente los científicos descubrieron que estos acontecimientos fueron causados por dos virus genéticamente distintos: el virus del Ébola de Zaire y el virus del Ébola de Sudán. Luego de este descubrimiento, los científicos concluyeron que el virus provenía de dos fuentes diferentes y se propagó de forma independiente a las personas en cada una de las áreas afectadas. [48]

La enfermedad generalmente ocurre en brotes en regiones tropicales del África subsahariana. Desde 1976 hasta 2018, la OMS notificó 31.095 casos con 12.950 muertes en general. El brote más grande hasta la fecha fue la epidemia del 2014-2016 en África occidental, que causó una gran cantidad de muertes en Guinea, Sierra Leona y Liberia.

La tasa promedio de letalidad de casos de EVE es de alrededor del 50 %. Las tasas de letalidad han variado del 25 % al 90 % en brotes pasados. La siguiente tabla, extraída de la OMS [58], muestra la cronología de los distintos brotes de la enfermedad del Ébola a través del tiempo.

| <b>Año</b>        | <b>País</b>             | <b>Ebolavirus</b> | <b>Casos</b> | <b>Muertes</b> | <b>% de letalidad</b> |
|-------------------|-------------------------|-------------------|--------------|----------------|-----------------------|
| 2018 - 2019       | R.D.C.                  | Zaire             | en marcha    |                |                       |
| 2018              | R.D.C.                  | Zaire             | 54           | 33             | 61 %                  |
| 2017              | R.D.C.                  | Zaire             | 8            | 4              | 50 %                  |
| 2015              | Italia                  | Zaire             | 1            | 0              | 0 %                   |
| 2014              | España                  | Zaire             | 1            | 0              | 0 %                   |
| 2014              | UK                      | Zaire             | 1            | 0              | 0 %                   |
| 2014              | USA                     | Zaire             | 4            | 1              | 25 %                  |
| 2014              | Senegal                 | Zaire             | 1            | 0              | 0 %                   |
| 2014              | Mali                    | Zaire             | 8            | 6              | 75 %                  |
| 2014              | Nigeria                 | Zaire             | 20           | 8              | 40 %                  |
| 2014 - 2016       | Sierra León             | Zaire             | 14124*       | 3956*          | 28 %                  |
| 2014 - 2016       | Liberia                 | Zaire             | 10675*       | 4809*          | 45 %                  |
| 2014 - 2016       | Guinea                  | Zaire             | 3811*        | 2543*          | 67 %                  |
| 2014              | R.D.C.                  |                   |              |                |                       |
| 2012              | R.D.C.                  | Bundibugyo        | 57           | 29             | 51 %                  |
| 2012              | Uganda                  | Sudán             | 7            | 4              | 57 %                  |
| 2012              | Uganda                  | Sudán             | 24           | 17             | 71 %                  |
| 2011              | Uganda                  | Sudán             | 1            | 1              | 100 %                 |
| 2008              | R.D.C.                  | Zaire             | 32           | 14             | 44 %                  |
| 2007              | Uganda                  | Bundibugyo        | 149          | 37             | 25 %                  |
| 2007              | R.D.C.                  | Zaire             | 264          | 187            | 71 %                  |
| 2005              | Congo                   | Zaire             | 12           | 10             | 83 %                  |
| 2004              | Sudán                   | Sudán             | 17           | 7              | 41 %                  |
| 2003<br>(Nov-Dic) | Congo                   | Zaire             | 35           | 29             | 83 %                  |
| 2003<br>(Ene-Abr) | Congo                   | Zaire             | 143          | 128            | 90 %                  |
| 2001 - 2002       | Congo                   | Zaire             | 59           | 44             | 75 %                  |
| 2001 - 2002       | Gabón                   | Zaire             | 65           | 53             | 82 %                  |
| 2000              | Uganda                  | Sudán             | 425          | 224            | 53 %                  |
| 1996              | Sudáfrica<br>(ex-Gabón) | Zaire             | 1            | 1              | 100 %                 |

| Año               | País          | Ebolavirus | Casos | Muertes | % de letalidad |
|-------------------|---------------|------------|-------|---------|----------------|
| 1996<br>(Jul-Dic) | Gabón         | Zaire      | 60    | 45      | 75 %           |
| 1996<br>(Ene-Abr) | Gabón         | Zaire      | 31    | 21      | 68 %           |
| 1995              | R.D.C.        | Zaire      | 315   | 254     | 81 %           |
| 1994              | Côte d'Ivoire | Taï Forest | 1     | 0       | 0 %            |
| 1994              | Gabón         | Zaire      | 52    | 31      | 60 %           |
| 1979              | Sudán         | Sudán      | 34    | 22      | 65 %           |
| 1977              | R.D.C.        | Zaire      | 1     | 1       | 100 %          |
| 1976              | Sudán         | Sudán      | 284   | 151     | 53 %           |
| 1976              | R.D.C.        | Zaire      | 318   | 280     | 88 %           |

\* Incluye casos de EVE sospechosos, probables y confirmados.

Figura 4.1: Cronología de brotes por EVE.

#### 4.1.1. Signos y síntomas.

Los síntomas pueden aparecer en cualquier momento desde 2 hasta 21 días posteriores al contacto con el virus, con un promedio de 8 a 10 días. Este período entre la exposición a una enfermedad y los síntomas se conoce como el período de incubación. Una persona infectada no puede contagiar a otras a menos que haya desarrollado sintomatología.

Los pacientes con Ébola generalmente experimentan fiebre, fatiga, dolor muscular y dolor de cabeza, seguidos de signos y síntomas variables que incluyen vómitos, diarrea, salpullido y diátesis hemorrágica que provocan sangrado externo, sangrado interno o ambos. En casos severos, puede desarrollarse una disfunción multiorgánica (por ejemplo, daño hepático, insuficiencia renal y afectación del sistema nervioso central), que conduce a shock y muerte. En la fase clínica inicial, el Ébola puede ser difícil de distinguir de otras enfermedades infecciosas, como la malaria, la fiebre tifoidea y la fiebre de Lassa.

La recuperación de la enfermedad depende de una buena atención clínica de apoyo y la respuesta inmune del paciente. La recuperación puede comenzar 7 a 14 días después del primer síntoma. La muerte del paciente a menudo ocurre debido a la disminución de la presión arterial provocada por sangrado severo y pérdida de fluidos corporales [20]. Los estudios demuestran que los sobrevivientes de la infección por el virus del Ébola tienen anticuerpos (proteínas producidas por el sistema inmunitario que identifican y neutralizan los virus invasores) que pueden detectarse en la sangre hasta 10 años después de la recuperación. Se cree que los sobrevivientes tienen cierta inmunidad protectora contra el tipo de Ébola que los enfermó.

### 4.1.2. Transmisión.

Se cree que los murciélagos frugívoros de la familia Pteropodidae son hospedadores naturales del virus del Ébola. La enfermedad se introduce en la población humana a través del contacto cercano con la sangre, las secreciones, los órganos u otros fluidos corporales de animales infectados (por ejemplo, murciélagos, chimpancés, gorilas, monos, antílopes del bosque o puerco espines) encontrados enfermos o muertos en la selva tropical. También si se prepara carne de animales infectados como comida. Esto se llama un evento indirecto. El virus luego se propaga a través de la transmisión de persona a persona mediante el contacto directo.

La transmisión del virus entre personas se produce a través del contacto directo con la sangre o los fluidos corporales (por ejemplo, orina, saliva, sudor, heces, vómitos, leche materna o semen) de personas sintomáticas o fallecidas o con objetos (por ejemplo, agujas y jeringas) contaminados con fluidos corporales de una persona infectada. El nivel de infectividad aumenta dramáticamente a medida que la enfermedad progresa y aumenta la carga viral de la persona infectada. Los fluidos, la piel y otros tejidos de las personas que mueren de Ébola son extremadamente infecciosos y representan un peligro para cualquier persona que tenga contacto sin protección con el cuerpo, incluidos los cuidadores y las personas que preparan el cuerpo para el entierro. [46, 49]

### 4.1.3. Tratamiento.

Los síntomas de la enfermedad se tratan a medida que aparecen. Cuando se aplican oportunamente las intervenciones básicas pueden mejorar significativamente las posibilidades de supervivencia. Éstas incluyen proporcionar líquidos y electrolitos a través de vía intravenosa, oxígeno-terapia para mantener el nivel de oxígeno, medicamentos para la presión arterial, los vómitos, la diarrea, la fiebre y el dolor, y el tratamiento de otras infecciones en el caso de que ocurran.

Actualmente se están desarrollando nuevas drogas antivirales para combatir la infección del virus del Ébola. Según la OMS y el CDC (Centro de Control de Enfermedades y Prevención) de Estados Unidos durante el brote del 2018-2019 en la R.D.C., se llevó a cabo el primer ensayo de control aleatorio de múltiples medicamentos para evaluar la efectividad y la seguridad de los medicamentos utilizados en el tratamiento de pacientes con Ébola bajo un marco ético desarrollado en consulta con expertos en el campo y la R.D.C. [58] Inicialmente, cuatro tratamientos de investigación estaban disponibles para tratar a pacientes con Ébola confirmado. Para dos de esos tratamientos, llamados regeneron (REGN-EB3) y mAb114, la supervivencia general fue considerablemente mayor. Estos dos medicamentos antivirales permanecen actualmente en uso para pacientes con Ébola confirmado. Los medicamentos que se están desarrollando para tratar la EVE funcionan al evitar que el virus haga copias de sí mismo. [47]

#### 4.1.4. Prevención y control.

Actualmente las vacunas para protección en contra del virus del Ébola se encuentran en proceso de desarrollo y han sido utilizadas para ayudar a controlar la propagación de los brotes de la enfermedad en Guinea y R.D.C. Según la OMS una vacuna experimental demostró ser altamente protectora contra EVE en un ensayo importante en Guinea en 2015. La vacuna, llamada rVSV-ZEBOV, se estudió en un ensayo que involucró a 11841 personas. Entre las 5837 personas que recibieron la vacuna, no se registraron casos de Ébola 10 días o más después de la vacunación. En comparación, hubo 23 casos 10 días o más después de la vacunación entre aquellos que no recibieron la vacuna. [58]

Por otro lado, la FDA de Estados Unidos (Food and Drug Administration) aprobó la vacuna contra el Ébola rVSV-ZEBOV el 19 de diciembre de 2019. La vacuna rVSV-ZEBOV es una dosis única que se ha encontrado segura y protectora contra sólo la especie Zaire de ebolavirus. Esta es la primera aprobación de la FDA de una vacuna contra el Ébola. Adicionalmente, se está desarrollando otra vacuna bajo un protocolo de investigación. Esta vacuna aprovecha dos componentes diferentes (Ad26.ZEBOV y MVA-BN-Filo) y requiere dos dosis con una dosis inicial seguida de una segunda dosis “de refuerzo” 56 días después. Esta vacuna también está diseñada para proteger contra la especie Zaire de ebolavirus. [45]

En general la OMS recomienda las siguientes medidas como protección y prevención para evitar la propagación de una epidemia de la enfermedad:

- Reducir el riesgo de transmisión de la vida silvestre al ser humano por el contacto con animales infectados (por ejemplo murciélagos, monos, simios, antílopes forestales o puerco espines) y el consumo de su carne cruda. Los animales deben manipularse con guantes y otra ropa protectora adecuada. Los productos animales (sangre y carne) deben cocinarse completamente antes de su consumo.
- Reducir el riesgo de transmisión de persona a persona por el contacto directo o cercano con personas con síntomas de EVE, particularmente con sus fluidos corporales. Se deben usar guantes y equipo de protección personal apropiado cuando se atiende a pacientes enfermos. Se requiere lavarse las manos regularmente después de visitar a los pacientes en el hospital, así como después de atender a los pacientes en el hogar.
- Medidas de contención de brotes, que incluyan el entierro seguro y digno de los muertos, la identificación de personas que pueden haber estado en contacto con alguien infectado con Ébola y el control de su salud durante 21 días, la importancia de separar a los sanos de los enfermos para evitar una mayor propagación, y la importancia de una buena higiene y mantener un ambiente limpio.

## 4.2. Modelando la epidemia.

En esta sección se describirá el modelo adoptado para predecir la dinámica y el control de la epidemia. El mismo se basa en un modelo estándar SEIR al cual se le

incorpora un compartimento adicional  $H$ , en el que se encuentran las personas hospitalizadas o en centros de control. Se asume que los agentes en esta clase se encuentran aislados del resto de la población por lo que no pueden contagiar la enfermedad. Al mismo tiempo, reciben tratamiento y en consecuencia su índice de supervivencia será mayor que el de la clase  $I$ . Estas muertes se reflejan en el modelo a través de la incorporación de un nuevo compartimento  $D$  en el que se contabilizan las muertes por la enfermedad. Uno podría argumentar que dicha clase es redundante ya que tanto los agentes que perecen así como los que se recuperan podrían moverse a la clase  $R$ . La sutileza de incorporar a la clase  $D$  es que se puede tener un seguimiento de las muertes que produce la epidemia e incorporar su valor en la función objetivo de los costos que se buscará minimizar. Adicionalmente se aprovecha el hecho de que las personas hospitalizadas reciben tratamiento y por esto la clase  $H$  se distingue de la clase  $Q$  del modelo de cuarentena. Por otro lado no se han considerado los nacimientos ni las muertes naturales ya que el análisis que se realizará estará basado en un corto período de tiempo y apuntado a controlar un brote epidémico. Se asume que la población total inicial  $N$  es lo suficientemente grande como para que no disminuya significativamente por la enfermedad, es decir,  $N - D \sim N$ .

El diagrama de flujo correspondiente es el siguiente:

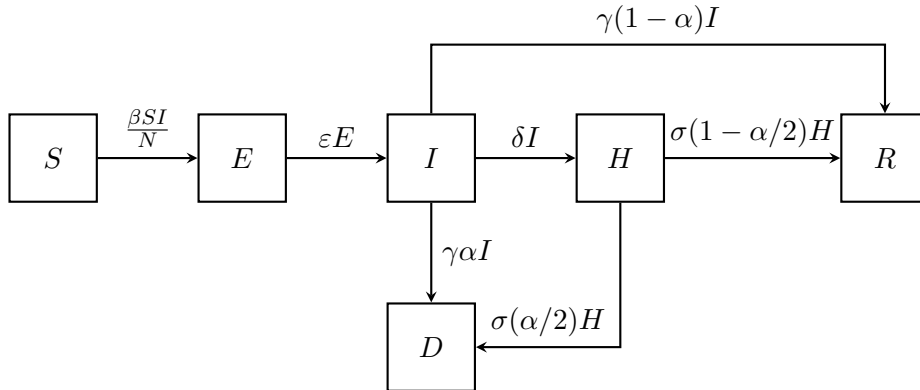


Figura 4.2: Diagrama de flujo, SEIHR con incidencia estándar y muertes por enfermedad.

A continuación algunas observaciones:

- Los individuos en la clase  $E$  se corresponden a aquellos que tuvieron un contacto adecuado con un infectado y se encuentran en el período de incubación o latencia. Permanecen en dicho compartimento hasta que presenten síntomas de la enfermedad y se vuelvan infecciosos, es decir, puedan transmitirla. Notar que el Ébola no se contagia a menos que se tenga un contacto con una persona que presenta sintomatología.  $\varepsilon$  es el coeficiente de latencia y  $1/\varepsilon$  es el período promedio de incubación, que en el caso del Ébola oscila entre 8 a 10 días.



- La incidencia usada es la estándar. El efecto de  $H$  es similar al de  $Q$  en el modelo de cuarentena, sin embargo, la incidencia elegida no es la ajustada por cuarentena. Se asume que  $N$  es lo suficientemente grande como para que la resta  $N - H \sim N$ .
- $\gamma$  es el coeficiente que representa el tiempo promedio de duración los síntomas de los individuos infectados que no se encuentran hospitalizados. En el caso de la enfermedad del Ébola se ha observado que la recuperación comienza alrededor de 7-14 días del primer síntoma.
- $\sigma$  corresponde al coeficiente que representa el tiempo promedio que un individuo debe permanecer hospitalizado y aislado. Se asume que como se encuentran en cuarentena el período de hospitalización debe ser igual o mayor a lo que dura la enfermedad en los agentes de la clase  $I$ . En otras palabras,  $1/\sigma \geq 1/\gamma$  y en consecuencia  $\gamma \geq \sigma$ . Se pondría pensar que los individuos hospitalizados se mantienen monitoreados y aislados hasta que se recuperan totalmente y dan negativo al test del virus.
- $\alpha$  representa a la fracción de individuos que mueren a causa de la enfermedad que no están hospitalizados. Se considera que el tratamiento recibido en los centros de salud u hospitales aumenta considerablemente la tasa de supervivencia. En este modelo en particular se ha tomado como  $\alpha/2$  la tasa de mortalidad por la enfermedad de los que se encuentran hospitalizados.
- $\delta$  representa la calidad del sistema de salud. En consecuencia tiene una correlación directa con la capacidad de la población para acceder a un centro de salud en el que se puede recibir los cuidados adecuados, es decir, aislamiento y tratamiento. Se asume que el primer día que una persona presenta síntomas buscará atención médica.  $\delta$  juega un papel fundamental en el modelo y depende del desarrollo socio económico de la población que se analiza. Idealmente los gobiernos deberían buscar que sea lo más grande posible, ya que esto los prepararía para enfrentar mejor una epidemia. Sin embargo, hay un costo asociado y depende principalmente de la inversión per cápita que destina cada país al sistema de salud. Tener una constante excesivamente grande podría representar también un desperdicio de recursos que nunca serán utilizados. Es necesario encontrar entonces un equilibrio ideal que asegure una salud óptima de la sociedad y que la prepare para resistir una epidemia sin desperdiciar recursos.

Finalmente las ecuaciones diferenciales asociadas al modelo serán las siguientes:

$$\begin{aligned}
S' &= -\frac{\beta SI}{N} \\
E' &= \frac{\beta SI}{N} - \varepsilon E \\
I' &= \varepsilon E - (\gamma + \delta)I \\
H' &= \delta I - \sigma H \\
R' &= \gamma(1 - \alpha)I + \sigma(1 - \alpha/2)H \\
D' &= \gamma\alpha I + \sigma(\alpha/2)H
\end{aligned} \tag{4.1}$$

Existe un equilibrio para este modelo, el de libre de enfermedad. Este equilibrio es aquel en el que no hay infectados, es decir,  $I = 0$ . Dado que no hay tasa de nacimientos en el sistema, este equilibrio resulta el trivial.

#### 4.2.1. Sistemas de salud en Guinea, Liberia y Sierra Leona.

El brote de Ébola del 2014-2015 fue el más grande y el peor de su tipo, poniendo en peligro a los frágiles sistemas de salud y la estabilidad económica de África Occidental. Aunque comenzó en diciembre de 2013 en una pequeña aldea en Guinea, no fue hasta el 8 de agosto de 2014 que la OMS anunció oficialmente al Ébola como emergencia de preocupación internacional. [38] La propagación de esta epidemia ha recibido diversas explicaciones, como por ejemplo, la movilidad humana entre las regiones afectadas, el comportamiento y las prácticas culturales como lo entierros tradicionales o consumo de animales. Sin embargo, uno de los factores más importantes son los ineficientes sistemas de salud de los países donde se presentaron la mayor cantidad de muertes por la enfermedad.

Guinea, Liberia y Sierra Leona han sufrido devastadoras guerras civiles las cuales deterioraron gravemente sus sistemas de salud e infraestructura. De acuerdo a las estadísticas de las Naciones Unidas [52] y la OMS [61], estos países tienen los índices de desarrollo humano más bajos con las infraestructuras de los sistemas de salud más débiles.

De acuerdo a [38] en Guinea, la densidad del personal sanitario (médicos, enfermeras, parteras, dentistas, farmacéuticos y psiquiatras) es menos de 1.5 por 10,000 habitantes, con un total de 3 camas de hospital por 10,000 habitantes. El gasto público en salud per cápita es de 9 dólares al año. Sierra Leona tiene una densidad de personal sanitario de 2.2 por 10,000 habitantes y aproximadamente 4 camas de hospital por 10,000 habitantes. Su gasto en salud per cápita es de 12 dólares. En Liberia, la densidad personal sanitario es menor a 3.7 por 10,000 habitantes, con alrededor de 8 camas de hospital por 10,000 habitantes y un gasto en salud per cápita de 13 dólares.

La salud pública y la infraestructura de los tres países carecía de los elementos esenciales necesarios para controlar un brote, incluyendo una fuerza laboral de salud fuerte. Sin embargo, el brote de Ébola tuvo también un impacto negativo significativo en el ya debilitado sistema de salud y contribuyó al decaimiento en la disponibilidad

de recursos humanos y materiales para la salud.

Según [23] en Liberia dos tercios de los residentes en zonas urbanas y la mitad de aquellos que vivían en zonas rurales afirmaron que fue bastante difícil o imposible acceder a cuidados de salud durante la epidemia de Ébola. En las zonas urbanas sólo el 20-30 % de los pacientes buscando cuidados durante la epidemia lograron conseguirlos mientras que en las zonas rurales este porcentaje aumenta a 70-80 %. Los pacientes que requerían cuidados prenatales y de obstetricia así como servicios de emergencia fueron aquellos que tuvieron mayor dificultad para acceder a dichos servicios.

En [38] se realiza un análisis detallado, según estadísticas y testimonios de expertos, en el que se explica las deficiencias de los sistemas de salud de los tres países durante la epidemia de Ébola. Se divide el análisis en seis categorías principales que conforman lo que la OMS considera los bloques esenciales de los sistemas de salud. Estos son *fuera laboral de salud, financiamiento de salud, información e investigación, productos y tecnologías médicas, liderazgo y gobernanza, y prestación de servicios*. Algunos puntos interesantes del artículo incluyen:

- La falta de trabajadores de salud calificados representó un gran obstáculo a la hora de contener el brote.
- La falta de inversión en entrenamiento de los trabajadores de la salud, infraestructura y la gestión de la cadena de suministro junto con el compromiso de la comunidad fue el mayor problema.
- Debido a la escasa vigilancia, la poca recolección de datos epidemiológicos y análisis estadístico, el brote no se detectó durante su fase inicial.
- Hubo retrasos en las pruebas y el diagnóstico debido a la escasez en la transferencia de muestras, el transporte, las ambulancias y los métodos de comunicación adecuados entre los funcionarios de salud, las aldeas y las zonas urbanas.
- La falta de liderazgo de los gobiernos a nivel nacional fue la razón principal que llevó a una pobre coordinación y una ausencia de pronta respuesta.

#### 4.2.2. Ajustando los parámetros.

Con el objetivo de que el modelo refleje características cuantitativas y cualitativas de la realidad se ajustaron los parámetros usando estadísticas obtenidas de la OMS y el CDC de los Estados Unidos [43,44]. Se ha tomado como unidad de tiempo a un día (24 horas), es decir que  $\Delta t = 1$  día. Teniendo esto en cuenta se han asumido los siguientes parámetros:

- Se asume que el tiempo de incubación promedio es de 8 días, es decir,  $\varepsilon = 1/8 = 0,125$ .
- Se considera que el tiempo promedio desde que una persona presenta el primer síntoma hasta que se recupera o muere es de 14 días, es decir,  $\gamma = 1/14 \sim 0,0714$ .

- El tiempo de hospitalización se toma en 15 días independientemente de cuando la persona presentó los primeros síntomas, es decir,  $\sigma = 1/15 \sim 0,0667$ . Notar que se cumple  $\sigma \leq \gamma$ .
- Se asume que la calidad del sistema de salud sólo permite que el 20% de los infectados se hospitalicen por unidad de tiempo. En consecuencia  $\delta = 0,2$ .
- La cantidad de población total se tomó igual a la reportada al inicio del 2014 por el Banco Mundial. [42]
- Se asume que el índice de mortalidad por la enfermedad es del 45%, es decir,  $\alpha = 0,45$ .

Se han utilizado los datos de los brotes en Guinea, Liberia y Sierra Leona quienes han sido los más afectados por la enfermedad y se han ajustado los parámetros del modelo en cada caso por separado. Los siguientes gráficos obtenidos de [44] demuestran la incidencia de los nuevos casos sospechados, probables y confirmados en cada país.

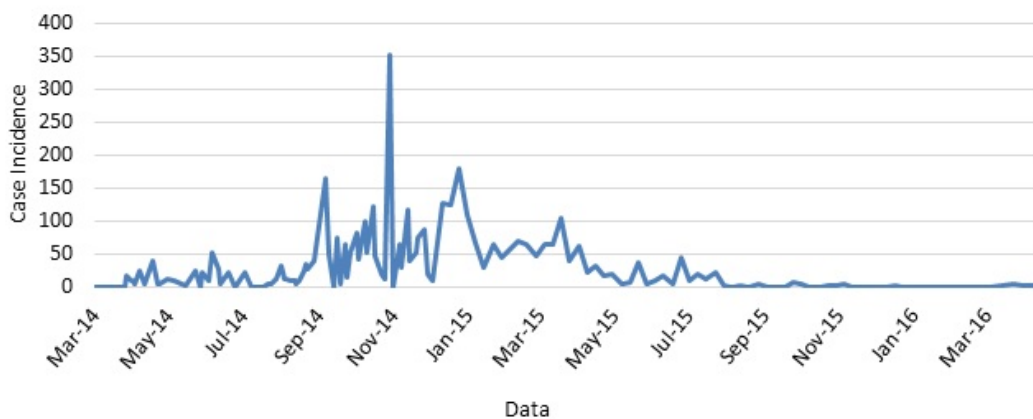


Figura 4.3: Incidencia de nuevos casos en Guinea durante el brote del 25 de marzo de 2014 al 13 de abril de 2016.

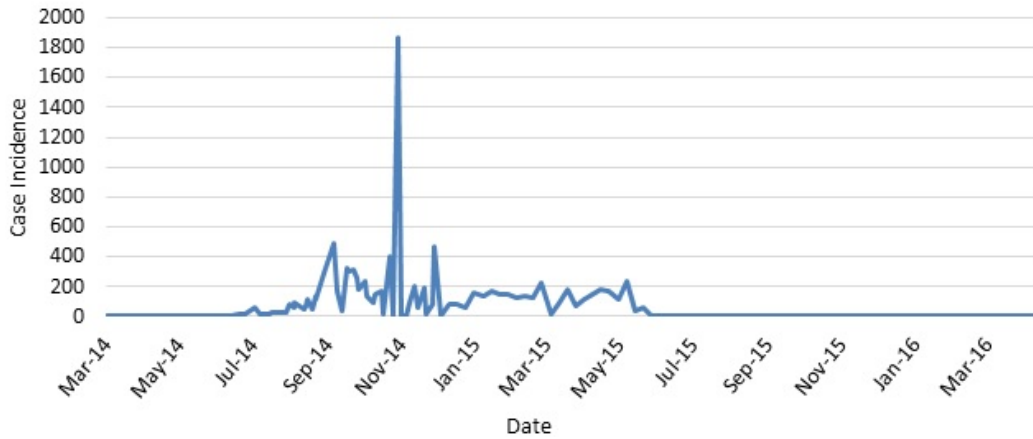


Figura 4.4: Incidencia de nuevos casos en Liberia durante el brote del 25 de marzo de 2014 al 13 de abril de 2016.

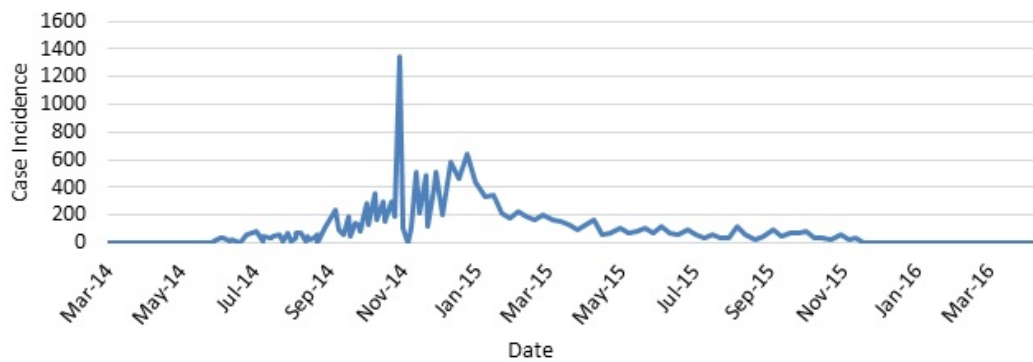


Figura 4.5: Incidencia de nuevos casos en Sierra Leona durante el brote del 25 de marzo de 2014 al 13 de abril de 2016.

En cada caso se han ajustado los parámetros  $\beta$  e  $I_0$  para que se minimice el error cuadrático medio entre la curva obtenida y los datos. Se han utilizado distintos períodos de tiempo para cada país, en los que se tuvieron en cuenta en qué momento se dio el pico de la enfermedad. Dado que se está ajustando el modelo no controlado es muy importante que se usen los datos correspondientes al brote antes de una gran intervención efectiva por parte de organismos internacionales para contener la situación. Por esta razón, sólo se han utilizado datos previos al máximo de incidencia. Las fechas implementadas han sido las siguientes: Guinea: 25 Marzo 2014 - 15 Octubre 2014. Liberia: 30 Abril 2014 - 3 Octubre 2014. Sierra Leona: 28 Mayo 2014 - 15 Octubre 2014. En los gráficos siguientes se muestran los datos de los casos acumulados suministrados por la OMS [43] y los casos acumulados por el modelo, es decir,  $I(t) + H(t) + R(t) + D(t)$ . Las condiciones iniciales fueron las siguientes:

- Guinea:  $S_0 = 12044$  (en miles),  $E_0 = 1$ ,  $I_0 = 36$ ,  $H_0 = 0$ ,  $R_0 = 0$ ,  $D_0 = 0$ .
- Liberia:  $S_0 = 4038$  (en miles),  $E_0 = 1$ ,  $I_0 = 6$ ,  $H_0 = 0$ ,  $R_0 = 0$ ,  $D_0 = 0$ .
- Sierra Leona:  $S_0 = 6939$  (en miles),  $E_0 = 1$ ,  $I_0 = 36$ ,  $H_0 = 0$ ,  $R_0 = 0$ ,  $D_0 = 0$ .

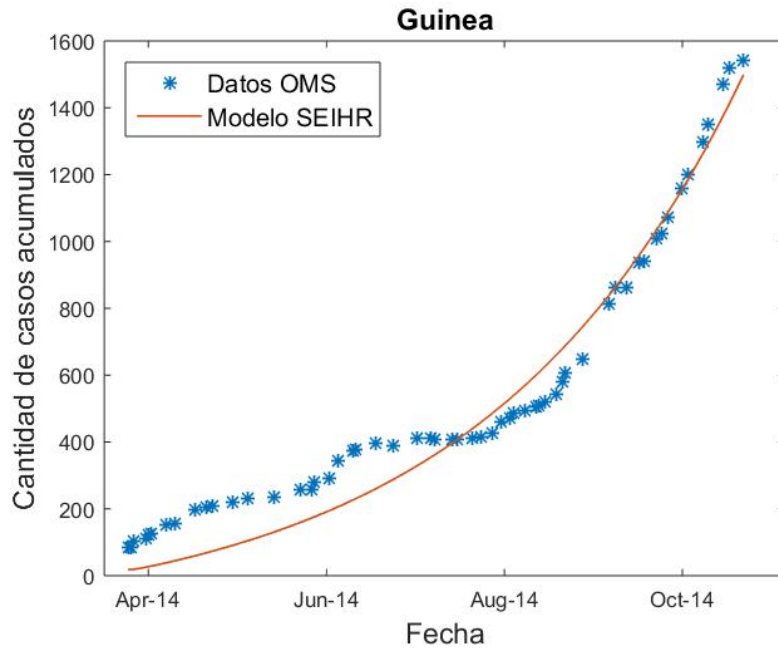


Figura 4.6: Comparación de casos acumulados - Guinea

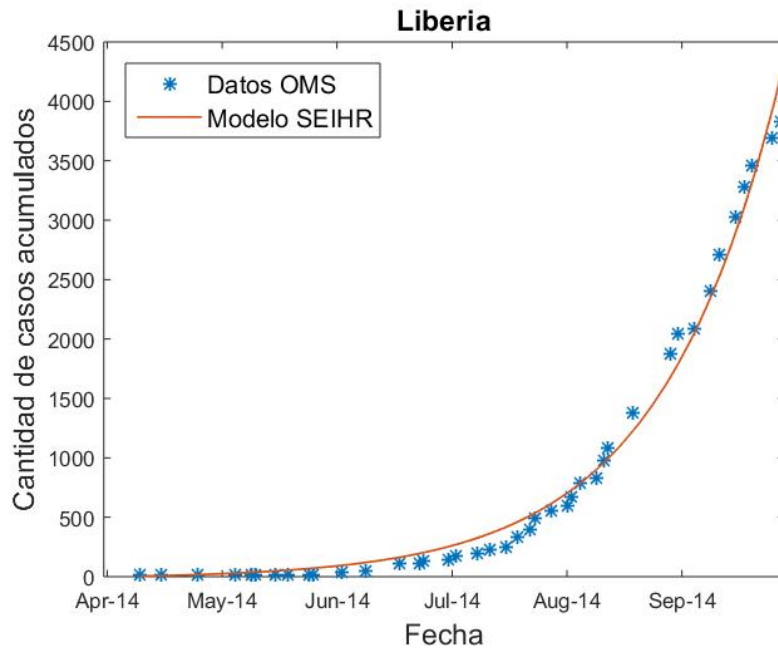


Figura 4.7: Comparación de casos acumulados - Liberia

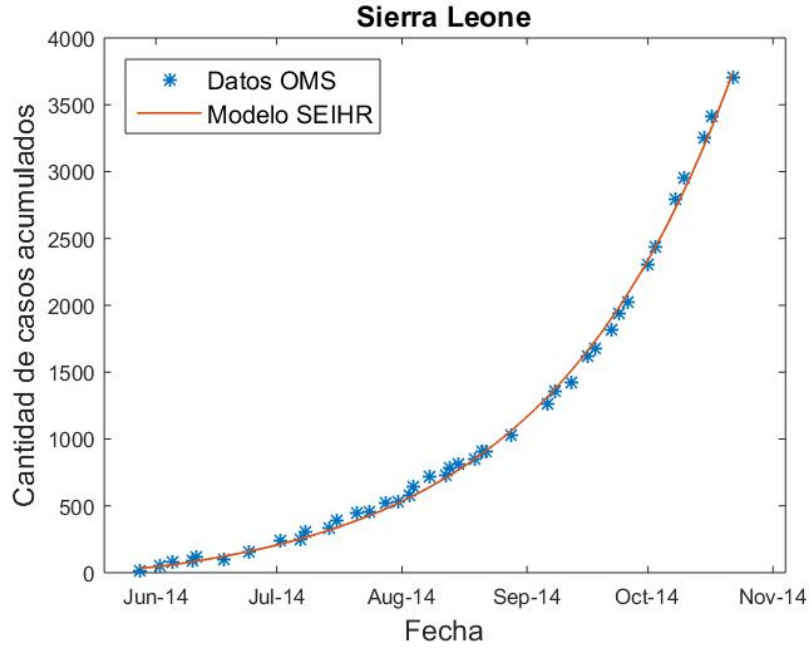


Figura 4.8: Comparación de casos acumulados - Sierra Leona

Una vez ajustados los parámetros del modelo se ha calculado también el número básico de reproducción  $R_0$ . Se ha utilizado la expresión  $R_0 = \frac{\beta}{\gamma + \delta}$ . Los resultados obtenidos para cada país se muestran a continuación en la siguiente tabla:

|              | $\beta$ | $R_0$  |
|--------------|---------|--------|
| Guinea       | 0.3078  | 1.1341 |
| Liberia      | 0.3887  | 1.4321 |
| Sierra Leona | 0.3428  | 1.2628 |

Figura 4.9: Parámetros  $\beta$  y  $R_0$  ajustados según datos de la OMS.

Dado que  $R_0 > 1$  se puede confirmar que la enfermedad provocará una epidemia y continuará extendiéndose en los tres países.

### 4.2.3. Definiendo el control y evaluando los costos.

Varios artículos han explorado diferentes controles óptimos en modelos epidemiológicos del Ébola [4, 29–33]. Algunos de ellos han estudiado el efecto de la vacunación y cómo regular la proporción de los vacunados por unidad de tiempo con el objetivo de evitar el brote de la enfermedad minimizando los costos asociados a la estrategia.

En este trabajo se propone explorar la estrategia del mejoramiento de los sistemas de salud. Como se ha descrito en la sección 4.2.1 los sistemas de salud de estos países se encontraban muy debilitados con el brote del Ébola.

LA OMS establece: “*Un sistema de salud que funciona adecuadamente trabajando en armonía es construido en base a tener trabajadores de salud entrenados y motivados, una infraestructura bien mantenida, y un razonable suministro de medicinas y tecnologías, respaldado con una inversión adecuada, fuertes planes de salud y una serie de políticas basadas en evidencias.*”

Estos pilares fueron seleccionados como un conjunto de prioridades de salud acordadas internacionalmente para crear una comprensión común de lo que es un sistema de salud y cómo puede fortalecerse. En concreto serían:

1. *Trabajadores de la salud.* Su disponibilidad debe distribuirse adecuadamente para ofrecer los mejores resultados a toda la población. El personal de salud debe estar bien calificado, capaz, sensible y eficiente.
2. *Financiamiento de salud.* El financiamiento de la salud es indispensable para mantener y mejorar el bienestar humano garantizando el empleo de la fuerza laboral, la disponibilidad de medicamentos y ofreciendo programas de promoción y prevención de salud pública.
3. *Información e investigación.* Un sistema de información e investigación que funcione bien asegura la recolección, análisis, distribución y comunicación efectiva y oportuna de la información.
4. *Productos médicos y tecnologías.* Los productos médicos (incluidos los medicamentos esenciales), las vacunas y las tecnologías deben estar disponibles y accesibles para la población. También deben ser de alta calidad y eficacia, y científicamente comprobados como seguros y rentables.
5. *Liderazgo y gobernanza.* Un sistema de salud bien dirigido y gobernado es uno que cuenta con marcos de políticas vitales, junto con una administración adecuada, asociaciones establecidas, el respeto de las regulaciones y la provisión de incentivos.
6. *Sistemas de distribución.* La buena prestación del servicio se logra cuando los servicios se entregan de manera oportuna, son rentables y seguros. Deben ser de alta calidad y de fácil acceso para toda la población, independientemente de su estatus social o ubicación geográfica. La prestación de servicios también incluye servicios centrados en la persona organizados en torno al paciente.

El estado de estos seis pilares se ve reflejado en el  $\delta$  del modelo. Este parámetro representa la calidad del sistema de salud del país y la posibilidad de las personas de acceder a él. El mismo puede ser mejorado con intervenciones del gobierno del país u organismos internacionales como la OMS.

Una vez establecido el control, el siguiente paso es buscar como medir los costos asociados a una determinada estrategia. En este trabajo se propone minimizar la cantidad de infectados ( $I$ ) a lo largo del tiempo así como también los costos de la inversión



directa ( $\delta$ ). Adicionalmente se desea minimizar las muertes causadas por la enfermedad al tiempo final  $T$ . Con ese objetivo en mente se propone el siguiente funcional:

$$J(\delta) = \int_0^T (I(t) + A \delta(t)^2) dt + B D(T)$$

Sujeto a la dinámica (4.1).  $A$  y  $B$  son pesos para establecer diferentes prioridades de la estrategia.

El problema de control óptimo sería entonces:

$$\min_{\delta \in \mathcal{A}} J(\delta) = \int_0^T (I(t) + A \delta(t)^2) dt + B D(T) \quad (4.2)$$

Donde la clase  $\mathcal{A}$  es aquella de las funciones continuas a trozos. Adicionalmente dado que los datos fueron ajustados con  $\delta = 0,2$  se toma dicho valor como cota inferior. Por otro lado tampoco sería real asumir que el 100% de la población pueda ser atendida. Se establece como cota superior al 80% lo que es equivalente a cuadruplicar la calidad del sistema existente (tarea no menor). En conclusión  $0,2 \leq \delta(t) \leq 0,8$ .

Siguiendo la línea de la teoría de control óptimo planteamos el Hamiltoniano del problema:

$$\begin{aligned} \mathbf{H} = & I + A \delta^2 + \lambda_S \left( -\frac{\beta SI}{N} \right) + \lambda_E \left( \frac{\beta SI}{N} - \varepsilon E \right) \\ & + \lambda_I (\varepsilon E - (\gamma + \delta)I) + \lambda_H (\delta I - \sigma H) \\ & + \lambda_R (\gamma(1 - \alpha)I + \sigma(1 - \alpha/2)H) + \lambda_D (\gamma\alpha I + \sigma(\alpha/2)H) \end{aligned}$$

Las ecuaciones diferenciales de  $\lambda$  son:

$$\begin{aligned} \lambda'_S &= \frac{\beta I}{N} (\lambda_S - \lambda_E) \\ \lambda'_E &= \varepsilon (\lambda_E - \lambda_I) \\ \lambda'_I &= -1 + (\lambda_S - \lambda_E) \frac{\beta S}{N} + \lambda_I (\gamma + \delta) \\ &\quad - \lambda_H \delta - \lambda_R \gamma (1 - \alpha) - \lambda_D \gamma \alpha \\ \lambda'_H &= \lambda_H \sigma - \lambda_R \sigma (1 - \alpha/2) - \lambda_D \sigma (\alpha/2) \\ \lambda'_R &= 0 \\ \lambda'_D &= 0 \end{aligned} \quad (4.3)$$

Con condiciones finales:

$$\begin{aligned} \lambda_S(T) &= 0 \\ \lambda_E(T) &= 0 \\ \lambda_I(T) &= 0 \\ \lambda_H(T) &= 0 \\ \lambda_R(T) &= 0 \\ \lambda_D(T) &= B \end{aligned} \quad (4.4)$$

Finalmente la condición de optimalidad:

$$\delta = \frac{I(\lambda_I - \lambda_H)}{2A} \quad (4.5)$$

### 4.3. Resultados obtenidos.

En esta sección se presentarán los resultados obtenidos de las simulaciones numéricas. Se tomó como unidad temporal ( $\Delta t$ ) a 1 (un) día y se utilizaron los parámetros ajustados para el modelo de Sierra Leona. El software utilizado fue Matlab junto con el paquete de optimización (Optimization Toolbox) para implementar el primer método numérico descrito en 3.4. Para esto, se ha usado el método de Runge Kutta de orden 4 y la regla de trapecios compuesta. Para la parte de optimización no lineal se ha usado la función `fmincon` de Matlab basada en el algoritmo del punto interior con método de barrera.

El primer escenario analizado ha sido el correspondiente al modelo ajustado para Sierra Leona en el día 15 de Octubre del 2014. Es decir, se dejó evolucionar el sistema sin variación del control ( $\delta \equiv 0,2$ ) hasta la fecha en cuestión. Tomando estas condiciones como iniciales ( $S_0 = 6939$  (en miles),  $E_0 = 661$ ,  $I_0 = 282$ ,  $H_0 = 635$ ,  $R_0 = 2052$ ,  $D_0 = 767$ ) se buscó el control óptimo a 100, 200 y 300 días variando el valor de la constante  $A$  mientras que se mantuvo el valor  $B = 1$ . La función objetivo a minimizar fue:

$$J(\delta) = \int_0^T (I(t) + A \delta^2) dt + B D(T)$$

Los resultados obtenidos fueron los siguientes:

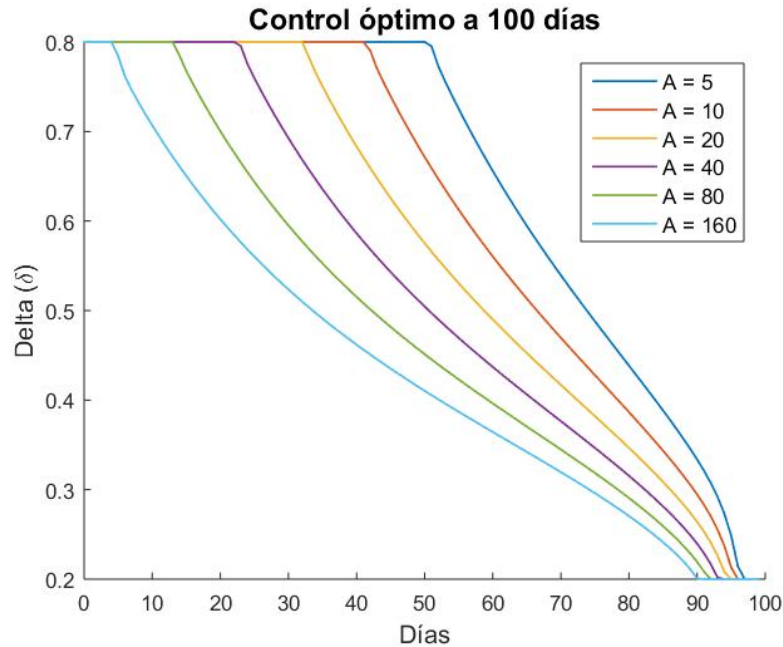


Figura 4.10: Controles óptimos a 100 días variando  $A$  y  $B = 1$ .

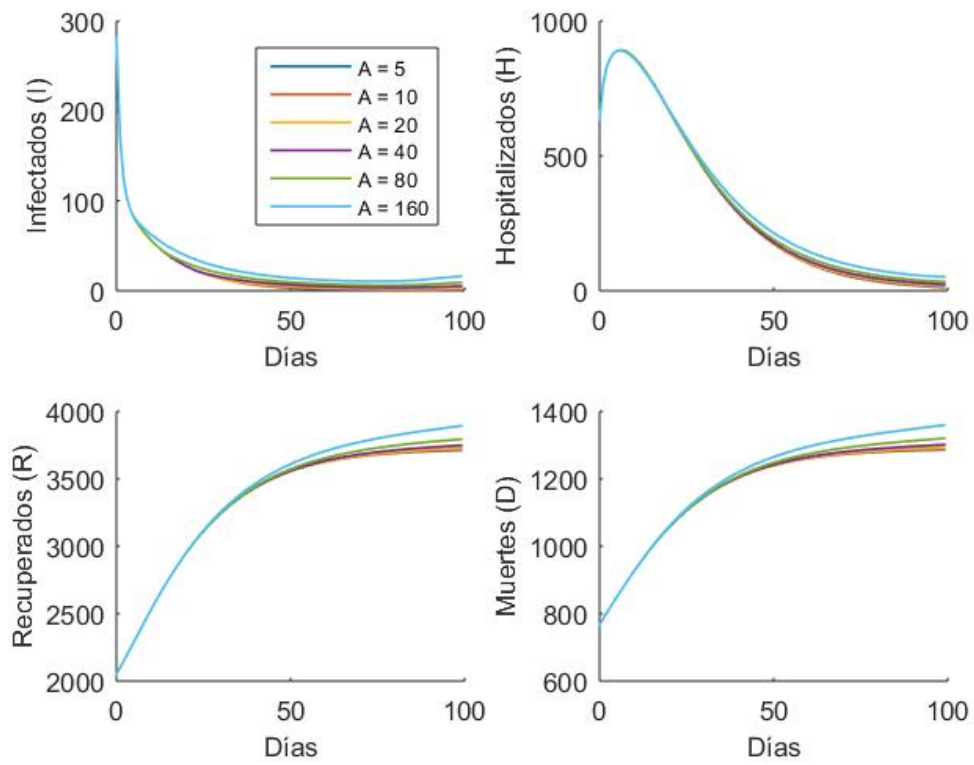


Figura 4.11: Dinámica del sistema asociada al control óptimo a 100 días.

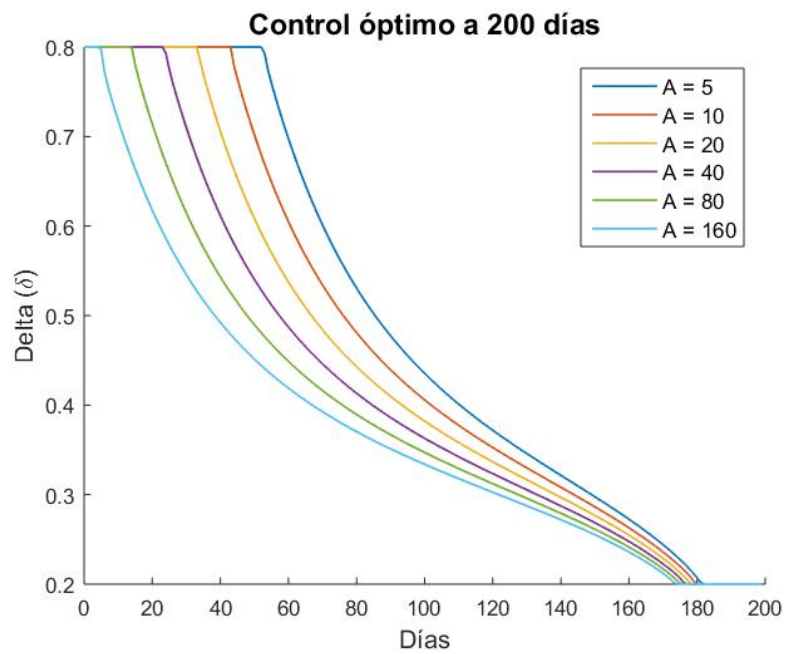


Figura 4.12: Controles óptimos a 200 días variando  $A$  y  $B = 1$ .

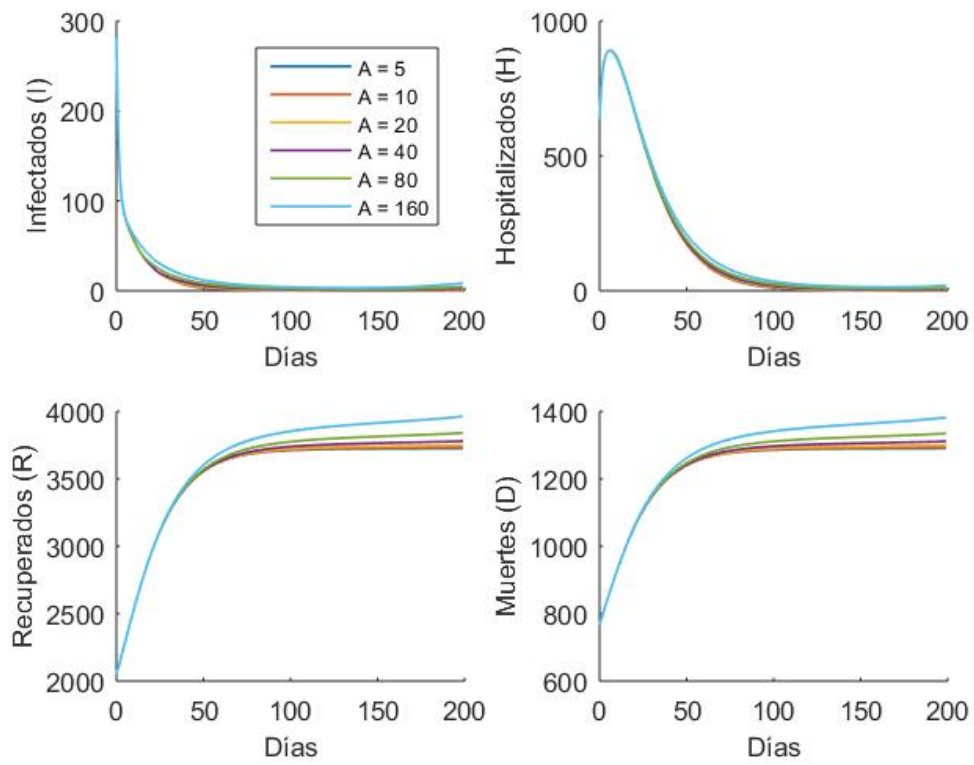


Figura 4.13: Dinámica del sistema asociada al control óptimo a 200 días.

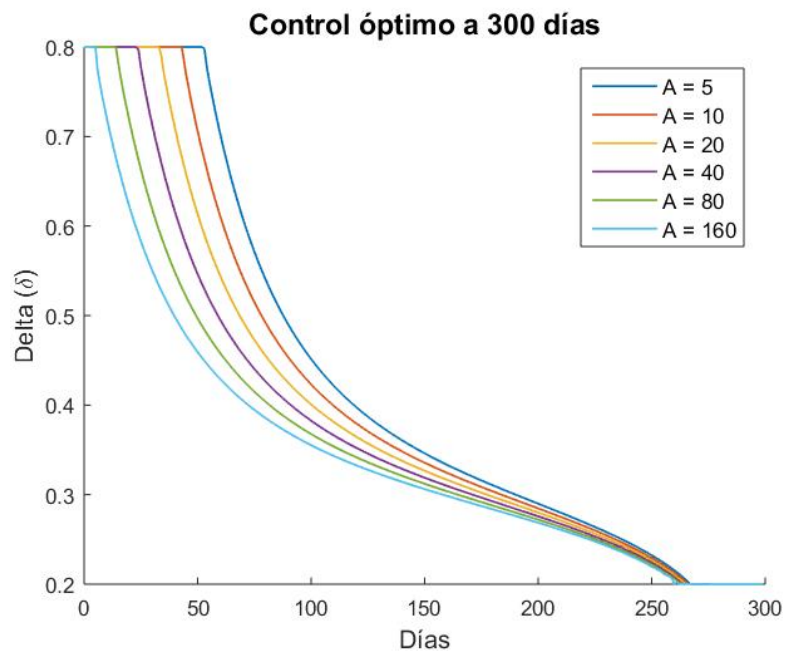


Figura 4.14: Controles óptimos a 300 días variando  $A$  y  $B = 1$ .

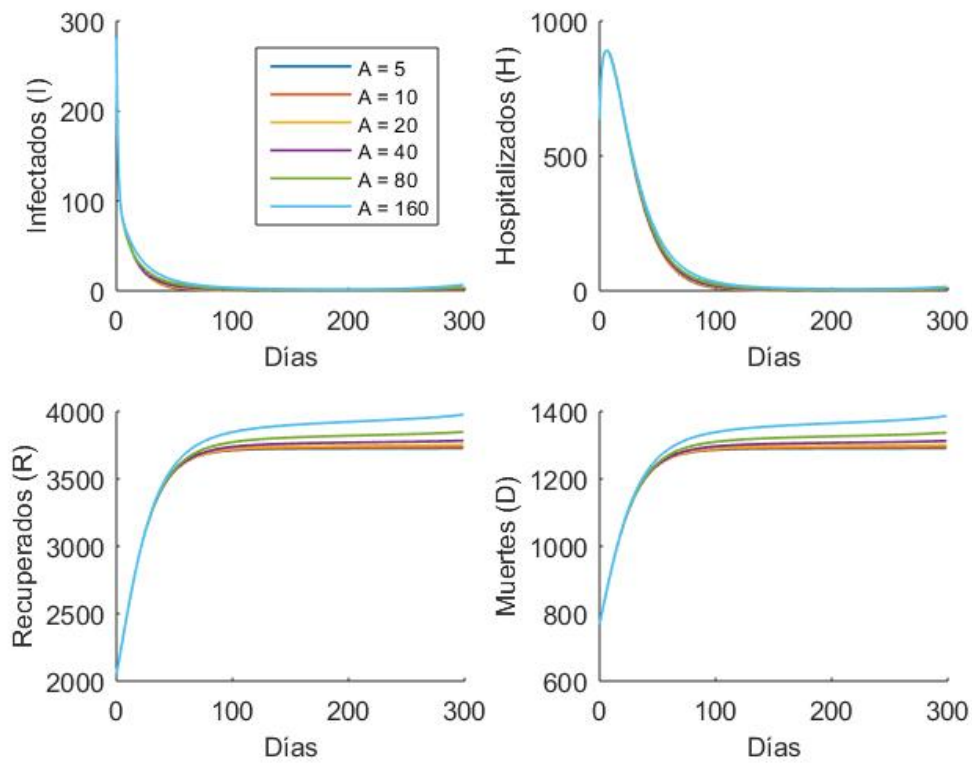


Figura 4.15: Dinámica del sistema asociada al control óptimo a 300 días.

El segundo escenario analizado ha sido el correspondiente al modelo ajustado para Sierra Leona desde el día cero (28 de Mayo de 2014). Es decir, se tomaron las mismas condiciones iniciales:  $S_0 = 6939$  (en miles),  $E_0 = 1$ ,  $I_0 = 36$ ,  $H_0 = 0$ ,  $R_0 = 0$ ,  $D_0 = 0$ . Se buscó el control óptimo a 200 días variando el valor de la constante  $A$  mientras que se mantuvo el valor  $B = 1$ . La función objetivo a minimizar fue:

$$J(\delta) = \int_0^T (I(t) + A \delta^2) dt + B D(T)$$

Los resultados obtenidos fueron los siguientes:

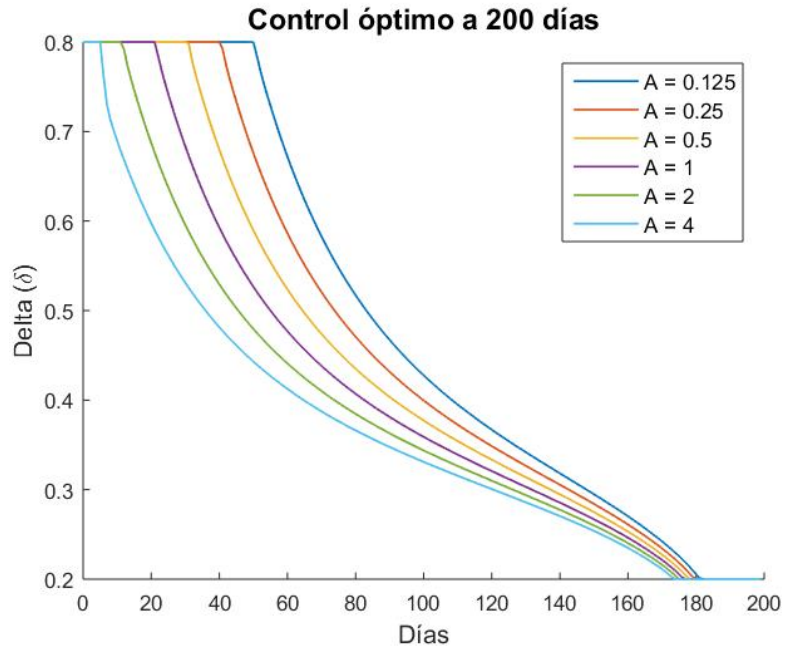


Figura 4.16: Controles óptimos a 200 días variando  $A$  y  $B = 1$ .

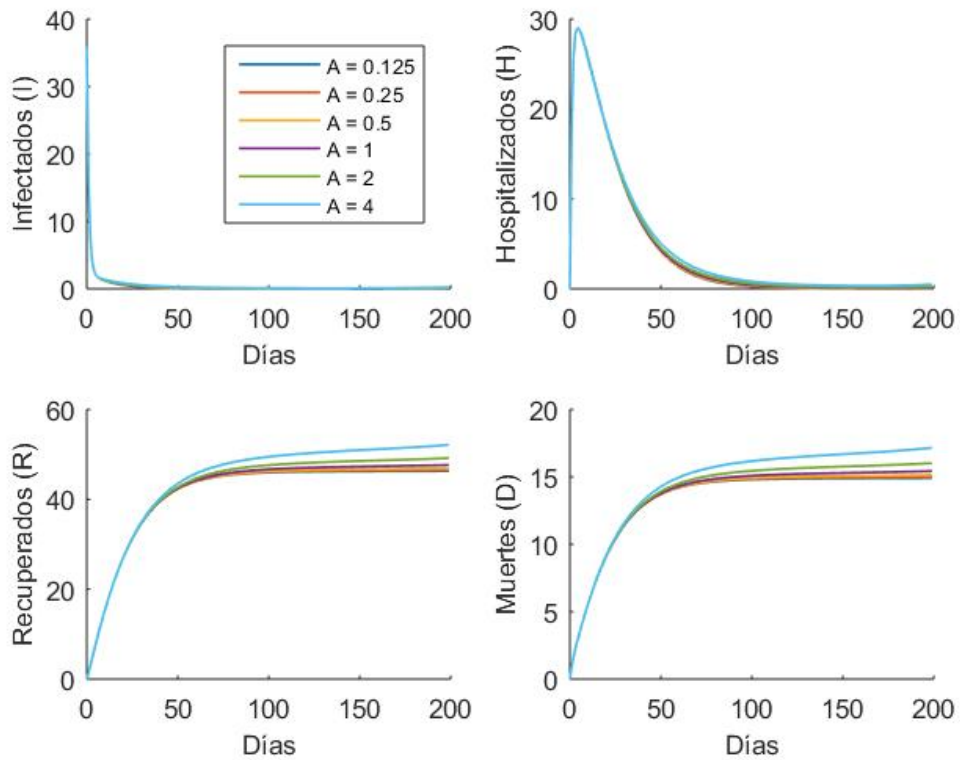


Figura 4.17: Dinámica del sistema asociada al control óptimo a 200 días.

El tercer y último escenario analizado ha sido el correspondiente al modelo ajustado para Sierra Leona desde el día cero (28 de Mayo de 2014). Es decir, se tomaron las mismas condiciones iniciales:  $S_0 = 6939$  (en miles),  $E_0 = 1$ ,  $I_0 = 36$ ,  $H_0 = 0$ ,  $R_0 = 0$ ,  $D_0 = 0$ . Se buscó el control óptimo a 200 días variando el valor de la constante  $B$  mientras que se mantuvo el valor  $A = 1$ . La función objetivo a minimizar fue:

$$J(\delta) = \int_0^T A \delta^2 dt + B \frac{D(T)}{R(T) + D(T)}$$

Los resultados obtenidos fueron los siguientes:

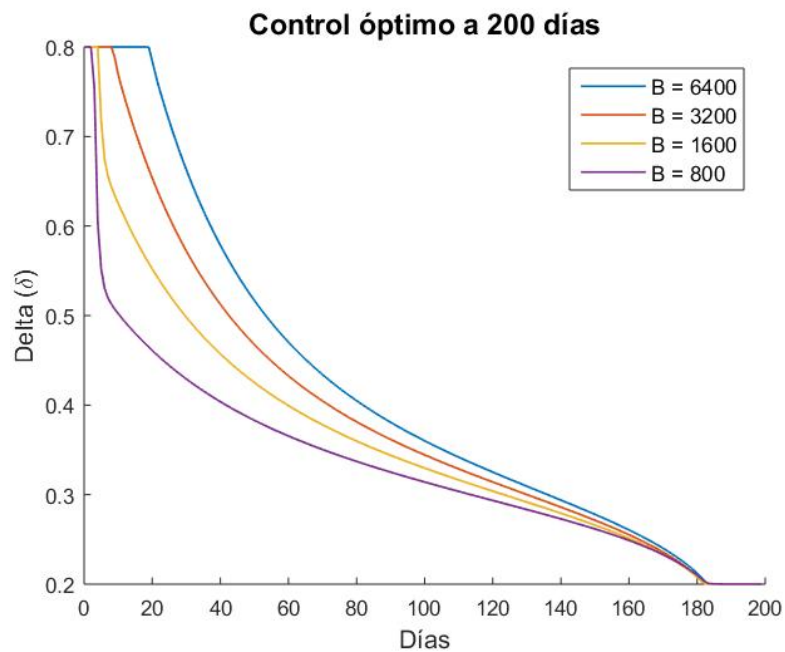


Figura 4.18: Controles óptimos a 200 días variando  $B$  y  $A = 1$ .

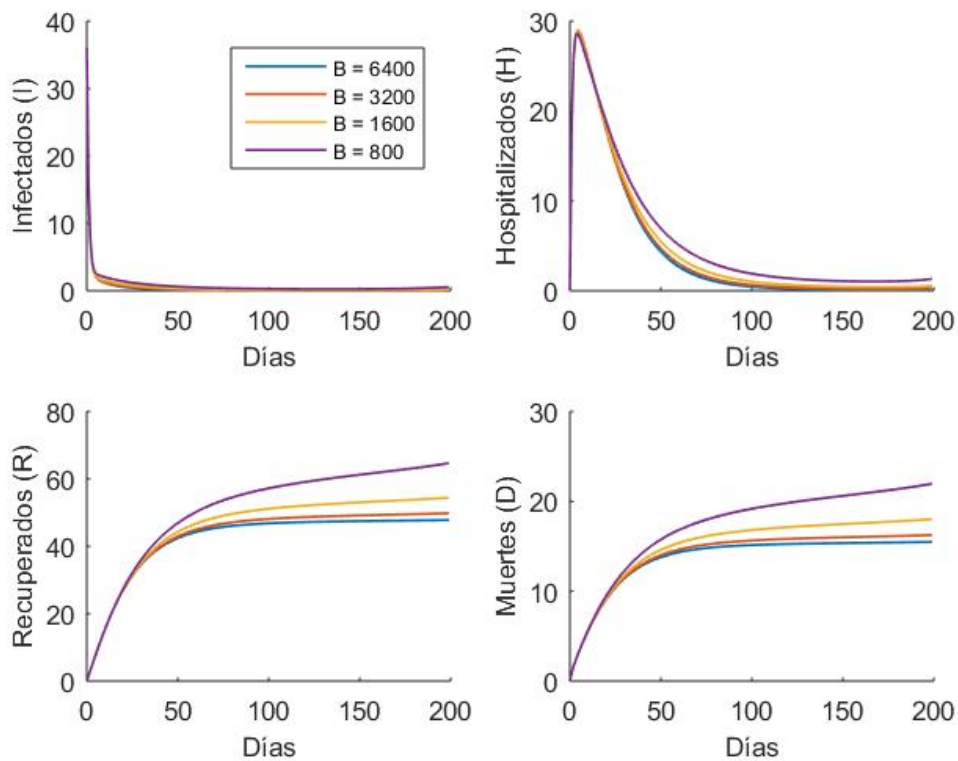


Figura 4.19: Dinámica del sistema asociada al control óptimo a 200 días.

#### 4.4. Discusión de los resultados.

En el primer escenario se buscó el control óptimo para minimizar la cantidad de infectados a lo largo de los días así como también la cantidad de muertes a tiempo final. Los pesos utilizados son arbitrarios ya que no se encontraron datos de costos reales cuantificables. En principio se eligieron estos valores para que todas las variables se encuentren en el mismo orden de magnitud.

Cuando se analizan las curvas a 100 días se ve que el control toma los valores máximos durante los primeros días y luego decrece de manera casi lineal hasta alcanzar la cota inferior al final del período. Independientemente de los valores de  $A$  todos los controles reflejan una estrategia agresiva al principio para reducir rápidamente la cantidad de infectados. En cuanto a la evolución del sistema se observa una mayor cantidad de muertes y recuperados al final del período para el control correspondiente a  $A = 160$  junto con una mayor cantidad de infectados.

Ahora si se analizan las curvas de 200 días se observa que también los controles toman los valores máximos al principio del intervalo para luego decrecer hacia la cota inferior al final. Sin embargo, en estas curvas el decrecimiento ya no sigue un patrón lineal. Una vez que la curva decrece lo hace rápidamente hasta aproximadamente el valor



0,4. Es interesante notar que las curvas parecen converger hacia el final del intervalo, lo cual puede deberse a que la cantidad de infectados se mantiene controlada mientras que la cantidad de hospitalizados se acerca a cero. Al igual que en las curvas de 100 días, las de 200 también reflejan una estrategia agresiva al principio. Si se observa la evolución del sistema se ve que las curvas de infectados y hospitalizados no presentan diferencias significativas mientras que sí se distingue una mayor cantidad de muertos y recuperados para el control con costo más alto ( $A = 160$ ).

En cuanto a las curvas de 300 días la única diferencia que se observa es una intensificación del patrón al que obedecen las de 200 días. Las curvas caen rápidamente hasta el día 100 alrededor del valor 0,4. La dinámica de la evolución del sistema es muy similar a la de 200 días sin diferencias significativas.

En el segundo escenario las condiciones iniciales son distintas. Hay una cantidad significativamente menor de infectados, expuestos y hospitalizados. Sin embargo, las curvas a 200 días del control son muy similares a las ya comentadas. La diferencia principal es que los valores de  $A$  son menores. Una explicación razonable de esto es que, por ejemplo, el costo para mejorar el sistema de salud en una ciudad es menor si se compara con el costo de mejorarlo en el país entero. Por lo tanto si la cantidad de infecciosos es baja se podría controlar la epidemia con costos menores apuntando la inversión y las estrategias de control hacia la región en donde se concentran los casos.

En el tercer escenario se modificó la función objetivo a minimizar. Se eliminó la variable  $I(t)$  dentro de la integral y se cambió el término a tiempo final. Lo que se busca aquí es minimizar la tasa de mortalidad de la enfermedad, definida como la cantidad total de muertes dividido la cantidad total de casos observados. Nuevamente los pesos se modificaron para que las variables estén en el mismo orden de magnitud. Es interesante notar que el patrón de las curvas del control son similares a las observadas anteriormente con la diferencia de que éstas convergen a un punto cerca del día 180. La curva correspondiente al menor peso presenta una caída significativa en los primeros 5 días desde el valor máximo hasta 0,5 aproximadamente.

Cuando se deja evolucionar al sistema con un delta constante igual a 0,2 (sin control) el índice de mortalidad es de 28,42%. En el primer escenario las curvas de 200 días muestran una mortalidad entre 24,36% – 24,71%. En el segundo escenario, el índice oscila entre 24,36% – 24,76%. Finalmente en el tercer escenario donde explícitamente se busca reducir al índice, la mortalidad se ubica en 24,36% – 24,57%. En resumen, no se observa mucha variabilidad en el índice de mortalidad en los diferentes escenarios hallándose todos en un rango de 4% menos que el sistema no controlado. Esta diferencia se explica gracias al hecho de aumentar la tasa de transferencia de la clase  $I$  a la clase  $H$  donde los agentes reciben tratamiento y donde mejoran sus probabilidades de sobrevivir a la enfermedad.

En conclusión, la característica común a todos los controles óptimos es una intervención agresiva en el principio. Esto se traduce en la realidad a mejorar significativamente el sistema de la salud. Es necesario crear centros de salud y de control provisorios distribuidos adecuadamente en la población afectada, donde los pacientes puedan recibir la atención necesaria o ser evaluados para su transferencia a unidades mejor preparadas.

Aumentar el número de trabajadores de salud entrenados correctamente y altamente calificados. Tener un sistema de vigilancia eficiente y adecuado con alertas tempranas y la capacidad de determinar las proyecciones de transmisión. Aumentar el financiamiento directo para mejorar la disponibilidad de suministros médicos, sistemas de transporte eficientes, métodos de comunicación y herramientas de diagnóstico. A pesar del apoyo internacional, se necesita inversión en la infraestructura de los sistemas de salud de los países para el desarrollo sostenible. La financiación debe incluir pagos adecuados, vigilancia de apoyo, compra de suministros necesarios y desarrollo de investigación. Finalmente, la mejora de las redes de rutas traslado de pacientes o muestras a tiempo y mejorar el acceso a la prestación de servicios de atención médica.

#### 4.5. Comentarios finales.

A lo largo de este trabajo se han presentado los modelos matemáticos básicos (y no tan básicos) de enfermedades infecciosas donde también se han analizado sus propiedades matemáticas. Se han definido los conceptos fundamentales como el número básico de reproducción y de reemplazo, la tasa de contactos adecuados e incidencia y la edad media de infección, entre otros. Luego se ha desarrollado la teoría de control óptimo exponiendo los resultados fundamentales así como los métodos numéricos para resolver los problemas. Finalmente se ha propuesto un modelo para el brote de Ébola del 2014 en los países de África occidental donde se analiza el papel fundamental que juegan los sistemas de salud. Se ha presentado un control para contener la epidemia y se lo ha simulado numéricamente exponiendo los resultados obtenidos.

En futuros trabajos sería interesante desarrollar nuevos modelos en los que se incorpore una saturación del sistema de salud y analizar cómo esto podría afectar la dinámica de la enfermedad. Por otro lado se podrían incorporar dimensiones espaciales al modelo mediante ecuaciones diferenciales parciales. La OMS ha publicado datos de las infecciones acumuladas por Ébola en varias ciudades de Guinea, Liberia y Sierra Leona.

El universo de los modelos matemáticos de enfermedades infecciosas es bastante extenso y también la bibliografía y los artículos científicos que los abordan. En la actualidad, el brote de la pandemia de Covid-19 ha puesto en jaque al mundo entero y posiblemente ha generado una de las peores recesiones económicas de la historia. Muchos países han solicitado a sus científicos desarrollar modelos que puedan brindar cierta claridad acerca de la evolución de la enfermedad y cómo adoptar las medidas óptimas para contener la situación generando el menor malestar posible en las sociedades. La importancia de estos trabajos es fundamental y tal vez sea, junto a los imprescindibles avances de la medicina, lo que nos ayude a superar los grandes desafíos que se han presentado.

# Apéndice A

## Métodos numéricos.

### A.1. Reglas de cuadraturas.

El objetivo de las reglas de cuadraturas es calcular aproximadamente el valor de una integral definida en un intervalo  $[a, b]$ . Con este fin supongamos que se desea saber el valor de  $\int_a^b f(t)dt$ . La idea básica es reemplazar a la función  $f$  por un polinomio  $p$  que esté “cerca” de la misma. Esto se debe a que, por un lado es fácil integrar polinomios y, por el otro, toda función continua puede aproximarse por estos.

Dados  $n+1$  puntos  $x_0, \dots, x_n \in [a, b]$  existe un único  $p_n \in P_n$  tal que  $p_n(x_j) = f(x_j)$ . Se define entonces la regla de cuadratura  $Q(f)$  por

$$Q(f) = \int_a^b p_n(t)dt.$$

Escribiendo a  $p_n$  en la forma de Lagrange, es decir,

$$p_n(x) = \sum_{j=0}^n f(x_j)\ell_j(x),$$

donde  $\ell_j(x_j) = 1$  y  $\ell_j(x_i) = 0$  si  $j \neq i$ , se tiene que

$$\int_a^b p_n(t)dt = \int_a^b \sum_{j=0}^n f(x_j)\ell_j(t)dt = \sum_{j=0}^n f(x_j) \int_a^b \ell_j(t)dt = \sum_{j=0}^n A_j f(x_j).$$

Los puntos  $x_j$  son llamados los nodos y los  $A_j$  los pesos de la integración numérica. Los pesos  $A_j = \int_a^b \ell_j(t)dt$  dependen sólo de los nodos  $x_j$  y una vez calculados se usan para aproximar la integral de cualquier función  $f$ .

A continuación se muestran las reglas de cuadraturas más comunes:

▪ Trapecios.

$$\int_a^b f(t)dt \sim T(f) = \frac{b-a}{2}(f(b) + f(a)).$$

El error de esta regla es  $-\frac{(b-a)^3}{12}f''(\xi)$  con  $\xi \in (a, b)$ .

- Simpson.

$$\int_a^b f(t)dt \sim S(f) = \frac{b-a}{6} \left( f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right)$$

El error de esta regla es  $-\frac{1}{90} \left(\frac{b-a}{2}\right)^5 f^{(4)}(\xi)$  con  $\xi \in (a, b)$ .

### A.1.1. Reglas de cuadraturas compuestas.

La idea es partir el intervalo  $[a, b]$  en subintervalos eligiendo puntos intermedios  $x_j$  con  $a = x_0 < x_1 < \dots < x_n = b$  y en cada uno de estos aplicar una aproximación del tipo Trapecios o Simpson. Se sabe que

$$I(f) = \int_a^b f(t)dt = \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} f(t)dt.$$

Luego si para cada

$$I_j(f) = \int_{x_j}^{x_{j+1}} f(t)dt$$

se tiene una fórmula de cuadratura  $Q_j(f)$  y se considera el error respectivo,

$$R_j(f) = I_j(f) - Q_j(f)$$

se obtiene que el error total es

$$R(f) = \int_a^b f(t)dt - \sum_{j=0}^{n-1} Q_j(f).$$

Por lo tanto la cuadratura será

$$\int_a^b f(t)dt \sim \sum_{j=0}^{n-1} Q_j(f) \quad \text{con error} \quad \sum_{j=0}^{n-1} R_j(f).$$

A continuación se muestran las reglas de cuadraturas compuestas más comunes:

- Trapecios compuesta. Supongamos que se divide el intervalo  $[a, b]$  en  $n$  subintervalos de igual longitud. Definimos entonces a  $h$  como la longitud de cada subintervalo, es decir,  $h = (b-a)/n$ . La regla de trapecios compuesta es

$$\int_a^b f(t)dt \sim T(f) = \frac{h}{2} \left[ f(a) + \sum_{j=1}^{n-1} 2f(a+jh) + f(b) \right]$$

El error de esta regla es  $-\frac{h^2}{12}(b-a)f''(\xi)$  con  $\xi \in (a, b)$ .

- Simpson compuesta. Al igual que antes se divide el intervalo  $[a, b]$  en  $n$  subintervalos de igual longitud. Como se interpola a  $f$  por un polinomio de grado 2, en puntos equiespaciados, en cada integral intervienen los nodos  $\left\{ x_j, \frac{x_j+x_{j+1}}{2}, x_{j+1} \right\}$ . Así el

paso  $h$  entre dos nodos de cada integral es la mitad de la longitud de  $[x_j, x_{j+1}]$ , es decir,  $h = \frac{1}{2} \frac{b-a}{n} = \frac{b-a}{2n}$ .

La regla de Simpson compuesta es

$$S(f) = \frac{h}{3} \left[ f(a) + 2 \sum_{j=1}^{n-1} f(x_j) + 4 \sum_{j=1}^{n-1} f\left(\frac{x_j + x_{j+1}}{2}\right) + f(b) \right]$$

El error de esta regla es  $-\frac{h^4}{180}(b-a)f^{(4)}(\xi)$  con  $\xi \in (a, b)$ .

## A.2. Métodos de Runge-Kutta.

Los métodos Runge-Kutta tienen como objetivo la resolución numérica de ecuaciones diferenciales ordinarias. Dada la ecuación diferencial  $x' = f(t, x)$  y un paso  $h$  se hace la aproximación de  $x(t+h)$  en base a  $x(t)$ . La idea es obtener aproximaciones que resulten del mismo orden que las que se obtienen con los métodos de Taylor, pero sin tener que calcular derivadas de  $f$ .

Por ejemplo si se considera un método de Taylor de orden 2 se tiene que

$$x(t+h) = x(t) + hT_2(t, x, h) + \mathcal{O}(h^3).$$

Se busca entonces  $\Phi_2(t, x, h)$  que no involucre las derivadas de  $f$  y que cumpla

$$x(t+h) = x(t) + h\Phi_2(t, x, h) + \mathcal{O}(h^3).$$

Para esto es suficiente que se cumpla

$$T_2(t, x, h) - \Phi_2(t, x, h) = \mathcal{O}(h^2).$$

Por lo tanto, se desea encontrar a  $\Phi_2$  de la forma

$$\Phi_2(t, x, h) = A_1 f(t, x) + A_2 f\left(t + \alpha h, x + \alpha h f(t, x)\right)$$

con  $A_1, A_2, \alpha$  a determinar. Observar que es natural que a un incremento  $\alpha h$  en la variable  $t$  le corresponda un incremento  $\alpha h x'$  en la variable  $x$ , siendo  $x' = f(t, x)$ .

A continuación se presentan los métodos de Euler modificado y Heun:

- Euler modificado.

$$x(t+h) \approx x(t) + \frac{h}{2} \left[ f\left(t + \frac{h}{2}, x(t) + \frac{h}{2} f(t, x(t))\right) \right]$$

- Heun.

$$x(t+h) \approx x(t) + \frac{h}{2} [f(t, x(t)) + f(t+h, x(t) + hf(t, x(t)))]$$

En forma análoga a lo hecho para orden 2, se pueden considerar más términos en el desarrollo de Taylor y deducir métodos de Runge-Kutta de mayor orden. En este trabajo utilizaremos el método de Runge-Kutta de orden 4 que se describe a continuación.

$$x(t+h) \approx x(t) + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

donde

$$\begin{aligned}k_1 &= f(t, x(t)) \\k_2 &= f\left(t + \frac{h}{2}, x(t) + \frac{h}{2}k_1\right) \\k_3 &= f\left(t + \frac{h}{2}, x(t) + \frac{h}{2}k_2\right) \\k_4 &= f(t+h, x(t) + hk_3).\end{aligned}$$

El error de este método es  $\mathcal{O}(h^4)$

# Bibliografía

- [1] R. M. Anderson, R. M. May - *Population biology of Infectious disease*. - Springer-Verlag, Berlin, Heidelberg, New York (1982).
- [2] R. M. Anderson, R. M. May - *The invasion, persistence, and spread of infectious diseases within animal and plant communities*. - Philos Trans R Soc Lond B 314: 533–570 (1986).
- [3] J. T. Betts - *Practical methods for optimal control and estimation using nonlinear programming*. - Advances in design and control, volume 19, Society for industrial & applied mathematics (2009).
- [4] H. Boujakjian - *Modeling the spread of Ebola with SEIR and optimal control*. SIAM Undergraduate Research Online. 9. 10.1137/16S015061. (2016)
- [5] R. Devaney, M. W. Hirsch - *Differential equations, dynamical systems, and an introduction to chaos*. - Pure and applied mathematics a series of monographs and textbooks 60, Academic Press (2004).
- [6] P. Drag, K. Styczen, M. Kwiatkowska, A. Szczurek *A review on the direct and indirect methods for solving optimal control problems with differential-algebraic constraints*. - Studies in computational intelligence, recent Advances in computational optimization, volume 610, (2016).
- [7] R. G. Duran, S. B. Lassalle, J. D. Rossi - *Elementos del cálculo numérico*. - Universidad de Buenos Aires, Facultad de ciencias exactas.
- [8] W. H. Fleming, R. W. Rishel - *Deterministic and stochastic optimal control*. - Applications of mathematics 1, Springer-Verlag, New York (1975).
- [9] A. Gulland - *West African countries plan to strengthen health systems after Ebola*. - BMJ, volume 349 (2014).
- [10] H. W. Hethcote - *The mathematics of infectious diseases*. - Society for Industrial and Applied Mathematics, volume 42, pages 599-653 (2000).
- [11] H. W. Hethcote - *Three basic epidemiological models*. - Applied mathematical ecology volume 18, pages 119—144 (1989).

- [12] H. W. Hethcote, M. Zhién, L. Shengbing - *Effects of quarantine in six endemic models for infectious diseases*. - Mathematical Biosciences, Elsevier Science, volume 180 (2002).
- [13] M. I. Kamien, N. L. Schwartz *Dynamic optimization: the calculus of variations and optimal control in economics and management*. - Advanced textbooks in economics, Elsevier science (1991).
- [14] M. J. Keeling, P. Rohani - *Modeling infectious diseases in humans and animals*. - Princeton University Press (2007).
- [15] M. P. Kieny, D. B. Evans, G. Schmets and S. Kadandalea - *Health-system resilience: reflections on the Ebola crisis in western Africa*. - Bulletin of the World Health Organisation, World Health Organisation, volume 92 (2014).
- [16] S. Lenhart, J. T. Workman - *Optimal control applied to biological models*. - Chapman & Hall/CRC mathematical and computational biology, CRC Press (2007).
- [17] A. Locatelli - *Optimal control: An introduction*. - Birkhäuser Basel (2001).
- [18] Z. Ma, J. Li - *Dynamical modeling and analysis of epidemics*. - World Scientific Publishing Company (2009).
- [19] S. Ma, Y. Xia - *Mathematical understanding of infectious disease dynamics*. - Lecture notes series, Institute for mathematical sciences, National University of Singapore, World Scientific Publishing Company (2008).
- [20] M. U. Majid, M. S. Tahir, Q. Ali, A. Q. Rao, B. Rashid, A. Ali, I. A. Nasir, T. Husnain - *Nature and history of Ebola virus: An overview*. - Archives of Neuroscience (2016).
- [21] M. Martcheva - *An introduction to mathematical epidemiology*. - Texts in applied mathematics volume 61, Springer (2015).
- [22] M. McAsey, L. Mou, W. Han - *Convergence of the forward-backward sweep method in optimal control*. - Computational optimization and applications, Springer US (2012).
- [23] P. A. McQuilkin, K. Udhayashankar, M. Niescierenko, L. Maranda - *Health-Care access during the Ebola virus epidemic in Liberia*. - The American Society of Tropical Medicine and Hygiene, The American Journal of Tropical Medicine and Hygiene, Volume 97, Issue 3, p. 931 - 936 (2017).
- [24] J.D. Murray - *Mathematical biology. I: An introduction*. - Interdisciplinary applied mathematics volume 17, Springer (2004).
- [25] S. Pedersen - *(Algunos) Modelos matemáticos para (algunas) enfermedades contagiosas: transmisión, infección, tratamiento*. - Tesis de licenciatura, Universidad de Buenos Aires, Facultad de ciencias exactas y naturales (2015).



- [26] L. Perko - *Differential equations and dynamical systems, third edition.* - Texts in applied mathematics, Springer (2006).
- [27] E. C. Pliego - *Modelos epidemiológicos de enfermedades virales infecciosas.* - Tesis de licenciatura, Benemérita Universidad Autónoma de Puebla (2011).
- [28] L.S. Pontryagin - *Mathematical theory of optimal processes.* - Classics of soviet mathematics, CRC Press (1987).
- [29] A. Rachah, D. Torres - *Analysis, simulation and optimal control of a SEIR model for Ebola virus with demographic effectes.* - Commun. Fac. Sci. Univ. Ank. Ser. A1 Math. Stat. 67 (2018), no. 1, 179–197.
- [30] A. Rachah, D. Torres - *Dynamics and optimal control of Ebola transmission.* - Mathematics in Computer Science, volume 10, issue 3, pages 331-342 (2016).
- [31] A. Rachah, D. Torres *Mathematical modelling, simulation, and optimal control of the 2014 Ebola outbreak in west Africa.* - Discrete Dynamics in Nature and Society, Hindawi Publishing Corporation, volume 2015, pages 1-9 (2015).
- [32] A. Rachah, D. Torres - *Modeling, dynamics and optimal control of Ebola virus spread.* - Pure and Applied Functional Analysis 1 (2016), no. 2, 277–289.
- [33] A. Rachah, D. Torres *Optimal control strategies for the spread of Ebola in West Africa.* - urnal of Mathematical Analysis, ISSN: 2217-3412, Volume 7, Issue 1 (2016).
- [34] A. V. Rao - *A survey for numerical methods for optimal control.* - Advances in the astronautical sciences (2010).
- [35] G. R. Rose - *Numerical methods for solving optimal control problems.* - University of Tennessee, Knoxville, Trace: Tennessee Research and Creative Exchange (2015).
- [36] R. Saleem, M. Habib, A. Manaf - *Review of forward backward sweep method for bounded and unbounded control problem with payoff term.* - Sci. Int. (2014).
- [37] S. Salsa, A. Squellati *Dynamical systems and optimal control.* - Bocconi Univ. (2018).
- [38] H. Shoman , E. Karafillakis, S. Rawaf - *The link between the West African Ebola outbreak and health systems in Guinea, Liberia and Sierra Leone: a systematic review.* - Globalization and Health, BioMed Central (2017).
- [39] T. C. Smith - *Ebola (deadly diseases and epidemics).* - Chelsea House Publishers (2006).
- [40] G. Teschl - *Ordinary differential equations and dynamical systems.* - AMS Graduate studies in mathematics 140, American Mathematical Society (2012).

- [41] W. Wang - *Backward bifurcation of an epidemic model with treatment*. - Mathematical Biosciences, Elsevier Science, volume 201 (2006).
- [42] Banco Mundial - *Países y economías*.  
<https://datos.bancomundial.org/pais>
- [43] Centers for Disease Control and Prevention - *2014-2016 Ebola outbreak in west Africa. Case counts*  
<https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/case-counts.html>
- [44] Centers for Disease Control and Prevention - *2014 Ebola outbreak in west Africa epidemic curves*.  
<https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/cumulative-cases-graphs.html>
- [45] Centers for Disease Control and Prevention - *Ebola Virus Disease, prevention and vaccine*.  
<https://www.cdc.gov/vhf/ebola/prevention/index.html>
- [46] Centers for Disease Control and Prevention - *Ebola Virus Disease, transmission*.  
<https://www.cdc.gov/vhf/ebola/transmission/index.html>
- [47] Centers for Disease Control and Prevention - *Ebola Virus Disease, treatment*.  
<https://www.cdc.gov/vhf/ebola/treatment/index.html>
- [48] Centers for Disease Control and Prevention - *History of Ebola Virus Disease*.  
<https://www.cdc.gov/vhf/ebola/history/summaries.html>
- [49] Centers for Disease Control and Prevention - *Overview, control strategies, and lessons learned in the CDC response to the 2014–2016 Ebola epidemic*.  
<https://www.cdc.gov/mmwr/volumes/65/su/su6503a2.htm>
- [50] National Geographic - *Sida*.  
<https://www.nationalgeographic.es/ciencia/sida>
- [51] Organización Mundial de la Salud - *La epidemia mundial de SIDA continúa creciendo*.  
<https://www.who.int/mediacentre/news/releases/2006/aids.epidemic.update/es/>
- [52] United Nations Development Programme - *Human Development Index and its components*.  
<http://hdr.undp.org/en/composite/HDI>
- [53] Wikipedia - *Peste negra*.  
[https://es.wikipedia.org/wiki/Peste\\_negra#La\\_enfermedad](https://es.wikipedia.org/wiki/Peste_negra#La_enfermedad)
- [54] Wikipedia - *Epidemia de ébola de 2014-2016*.  
[https://es.wikipedia.org/wiki/Epidemia\\_de\\_ébola\\_de\\_2014-2016](https://es.wikipedia.org/wiki/Epidemia_de_ébola_de_2014-2016)

- [55] Wikipedia - *Gran peste de Londres*.  
[https://es.wikipedia.org/wiki/Gran\\_peste\\_de\\_Londres](https://es.wikipedia.org/wiki/Gran_peste_de_Londres)
- [56] Wikipedia - *List of epidemics*.  
[https://en.wikipedia.org/wiki/List\\_of\\_epidemics](https://en.wikipedia.org/wiki/List_of_epidemics)
- [57] Wikipedia - *Peste Antonina*.  
[https://es.wikipedia.org/wiki/Peste\\_antonina](https://es.wikipedia.org/wiki/Peste_antonina)
- [58] World Health Organisation - *Ebola virus disease*.  
<https://www.who.int/en/news-room/fact-sheets/detail/ebola-virus-disease>
- [59] World Health Organisation - *Global Health Observatory (GHO) data - World Health Statistics 2014*.  
[https://www.who.int/gho/publications/world\\_health\\_statistics/2014/en/](https://www.who.int/gho/publications/world_health_statistics/2014/en/)
- [60] World Health Organisation - *Global Health Observatory (GHO) data - World Health Statistics 2015*.  
[https://www.who.int/gho/publications/world\\_health\\_statistics/2015/en/](https://www.who.int/gho/publications/world_health_statistics/2015/en/)
- [61] World Health Organisation - *Global Observatory on eHealth - key indicators*.  
<https://apps.who.int/gho/data/node.goe.keyind?lang=en>
- [62] World Health Organisation - *Update and summary guide to the report: Advancing the right to health – The vital role of law*.  
[https://www.who.int/healthsystems/topics/health-law/ARH-Summanry-Guide\\_Final\\_online\\_version.pdf?ua=1](https://www.who.int/healthsystems/topics/health-law/ARH-Summanry-Guide_Final_online_version.pdf?ua=1)