Modeling and Analysis of Mood Dynamics*

Hugo Gonzalez Villasanti and Kevin M. Passino
Dept. Electrical and Computer Engineering
Mathematical Biosciences Institute
The Ohio State University, 2015 Neil Ave., Columbus, OH 43210
August, 2019

Abstract

A nonlinear ordinary differential equation model of mood dynamics for illnesses on the bipolar spectrum is introduced. While this model ignores some key aspects of the illnesses, it does represent progress as it represents almost all those features considered by past models and more, and it is the first to include a range of cross-cutting features—while establishing the foundation for a general analysis of dynamics of mood disorders. In particular, we perform an analysis of conditions for mood to stabilize to euthymia (that have psychotherapeutic implications), and for pharmacotherapy to result in a switch from depression to mania when bipolar disorder was misdiagnosed as major depressive disorder and an anti-depressant is administered without a mood stabilizer.

1 Introduction

Mood disorders are prevalent and disabling illnesses that have received significant attention from clinical (e.g., psycho/pharmaco therapy), scientific (e.g., genetic or neurotransmitters), and technological (e.g., ECT or rTMS) perspectives [1–3]. Foundational to this work are mathematics (e.g., statistics or electromagnetic theory for the axon) and engineering (e.g., electrical engineering for rTMS/ECT devices). Here, we use a mathematical (nonlinear) dynamical system model to integrate scientific findings about mood disorders, predict features of mood via nonlinear and computational analyses, and connect these to psychotherapeutic practice. It is natural to take a dynamical systems approach to psychodynamics, especially for mood disorders, considering, for example, mood “swings” in bipolar disorders or recurrent depression. Our approach is a type of “meta-analysis” in that we do not conduct new experiments, but use results of multiple existing experimental studies, and link these together to obtain a dynamical systems-level representation and conclusions. We build on, then move past, the reductionistic paradigm.

*gonzalezvillasanti.1@osu.edu; passino.1@osu.edu
The mathematical and computational modeling and analysis of mood has been studied using a number of different approaches. In [4] the bipolarity of mood is characterized using a thermodynamics perspective, and fixed, periodic, and chaotic attractors are discussed; however, no mathematical models are used. Such dynamics, however, are partially supported by the self-report data from bipolar patients analyzed in [5], which shows that mood swings are not truly cyclic, but chaotic. Studies that employed specifically-designed experiments to validate their models are mostly limited to drift-diffusion models on binary choice (“Flanker task”), studying the role of rumination, attention, and executive functions in mood disorders [6, 7]. However, these studies are constrained to only specific features of mood. On the other hand, stochastic nonlinear models have been constructed to represent several dynamic features of mood disorders, with different levels of connection to neurobiological and psychological determinants. The effects of noise on bifurcations and episode sensitization in mood disorders have been described with linear oscillator models [8,9]. In [10,11] bipolarity is envisioned as arising from two possible types of regulation of a bistable system that result in mood oscillations characterized by two variables, depression and mania, with some connections to mood disorder determinants and therapy. Analysis in [12] considered a nonlinear limit cycle model of mood variations based on biochemical reaction equations. In [13] a model of the behavioral activation system, linked to bipolar disorder episodes, is constructed via a nonlinear stochastic model, and mono- and bi-stability of mood oscillations are studied. Also, the effects of noise and changes in nonlinearities are studied by comparing simulations with empirical and observational data. Mood regulation is analyzed in [14] with an inverted pendulum model, and therapy interventions represented as feedback controllers. The mixed state is accounted for in the oscillator model in [15]. While these studies focus on a particular subset of depressive disorder determinants using up to two state variables, integrative dynamical models accounting for biological and psychological determinants of depressive disorder were proposed in [16], with an emphasis on psychosocial states, and in [17], with an emphasis on neurobiological factors. Additional models considering depression are in [18] where a finite-state machine is used, and [19] where a stochastic model of random aspects of mood is employed. Also, there are a number of general computational/mathematical models used in psychiatry and psychology (e.g., see [6, 20, 21]) that are relevant to the above models. All the above studies, however, are limited in terms of experimental model validation, computational and mathematical analysis, and results obtained from analysis.

Bipolar I disorder (BD-I) and major depressive disorder (MDD) have large mood variations and clear relations to other mood disorders; hence, these two will take a prominent role in this work. Mood dynamics of bipolar II disorder (BD-II) and cyclothymia are considered here to be a special case of those of BD-I, and persistent depressive disorder (dysthymia) mood dynamics, a special case of MDD (in range of mood variation, not necessarily depressed mood persistence). This is partly justified by the close relations between mood symptoms of these disorders in DSM-5 (e.g., compare BD-I to BD-II or cyclothymia, BD-I to MDD) [22]. Yet, mood disorders on the “bipolar spectrum” [1] have different symptoms, severity, and experienced-time-length requirements for diagnosis (e.g., consider the differences between unipolar and bipolar depression [23]). Also, there are some symptoms that are significantly impairing and cannot be ignored by a clinician, but are essentially ignored here (e.g., expansive/racing thoughts and risk-taking in mania, suicidal ideation in depression, or sleep/weight disruption for several mood disorders). Colloquially, and in some scientific/clinical litera-
ture, mania and depression are sometimes described as opposite “poles” of a single vertical axis with mania at the top (“up”) and depression at the bottom (“down”) (but, see [24]). Yet, in the bipolar “mixed state” [25] there are both manic and depressed symptoms at the same time [22]. The mixed state is relatively common [1]. Also, in [26] the authors say that (i) “Goodwin and Jamison (1990) found that symptoms of depression and irritability, not simply elation, occur in 70%-80% of patients with mania” (see [1]) and (ii) “Goldberg et al. (2009), in findings from the STEP-BD study (n=1380), found a majority of bipolar patients with a full depressive episode have clinically relevant manic symptoms” (see [27]). Furthermore, see also pp. 144-147 in [28] where dimensional and categorical approaches to diagnosis based on depression and mania variables are used, with the mixed state corresponding to different combinations of severity levels for depression and mania (Figure 8.1 on p. 145 of [28] is related to Figure 1d below). Also, the existence of a mixed (mood) state is logically consistent with the literature on affective dimensions that identify mixed affect/emotion [29]. The mixed state cannot be understood in terms of a single axis with a variable that indicates whether someone is up or down; here, we include a second axis/variable so that there is one for mania severity, and another for depression severity. This does not imply that only two variables need to be used in order to represent mood dynamics (or symptoms related to mood). Indeed, below, there will be many other variables that influence the evolution of the mania and depression variables, and the dynamics of these.

In this paper, unlike the above past research mentioned above, the mathematical model represents a wider range of features: (i) euthymia, mania, depression, the mixed state, anhedonia, hedonia, and flat or blunted affect (all on a continuum of numeric values that represent the extent to which someone possesses a mood characteristic, or its “severity level,” as in “this person is fully manic, is close to euthymia, or is severely depressed”); (ii) mood dynamics for smooth variations between mood states (e.g., mania to euthymia to depression, and back, and severity levels for these as mood changes continuously); and (iii) “attractors” (“traps”) that mood can fall into and get stuck (e.g., being stuck in a severe depressed state), ones that are parameterized in terms of a person’s diagnosis (e.g., BD-I or MDD), and other characteristics of a person (defined below); some justification for such an attractors (basins of attraction) approach is given in [30] where such basins were found experimentally for a group depressed of patients. Also, a computational analysis of mood attractors is given for multiple cases on the bipolar spectrum (e.g., for BD-I, BD-II, cyclothymia, and MDD, and for each, with attractors for euthymia, depression, mania, and the mixed state). This shows how, in a sense, the characterization of mood disorders here can be related to a “dimensional diagnosis” (i.e., when mood stays in certain regions, or visits a region for some length of time, a certain type of disorder is indicated). Also unlike the above work, our mathematical model and analysis provides conditions for when a person will return to euthymia no matter how their mood is perturbed (i.e., when euthymia is a “global” mood attractor). It is explained how these conditions have clinical implications, in psycho and pharmacotherapy, for how to stabilize a person to euthymia. Also, our analysis provides a novel explanation of the mechanism underlying the mood stabilizer (e.g., Lithium Carbonate, \( \text{Li}_2\text{CO}_3 \)), that illustrates how it changes from being an anti-manic to anti-depressant agent, and also how it operates in the mixed state, serving simultaneously as an anti-manic and anti-depressant agent. These results resolve the statement where researchers identify the “...paradoxical effects of Lithium as both an antidepressant and antimanic agent” [1]. Next,
it is shown that for some depressed persons, if they are given an anti-depressant, the medicine could result in a mood trajectory that moves to the fully manic state (e.g., someone who was diagnosed with MDD, was not on a mood-stabilizer, and was driven fully manic by the anti-depressant—indicating that they might actually have BD-I).

To the limit scope of our work, we ignore: (i) details of emotion regulation and “fast” dynamics of emotions that occur on a time scale of less than 2-5 seconds [31] (but do consider the longer-term influence of emotions on mood as in [32]); (ii) effects of stress (e.g., see [33]); (iii) mood-congruent attention (see, e.g., [1, 32, 34]); (iv) positive/negative rumination and inter-episode features (e.g., see [35, 36]); (v) behavioral activation/inhibition system (BAS/BIS) sensitivity (e.g., see [37, 38]); (vi) goal pursuit (e.g., [39]) and goal dysregulation [40]; and (vii) “slow” dynamics that occur on a time scale of multiple years (e.g., seasonal influences [1, 41]). Also, we largely do not include the neural level, but the model nonlinearities are informed by it via referenced scientific studies.

The objective here is to uncover principles of mood dynamics that are common across the spectrum, and only consider other features in how they influence mood, and not via an explicit model of their own dynamics (e.g., a dynamical model of emotion regulation dysfunction or the coupling between mania and sleep). The ability to consider cross-spectrum issues arises here, in part, due to the use of a “dimensional approach,” as opposed to the “categorical” one in [22] (a comparative analysis these two approaches for modeling work is beyond the scope of this paper).

2 Mathematical Modeling of Mood Dynamics

2.1 Mood States, Trajectories, and Regions

Mood is represented in two dimensions here; however, consider briefly the one-dimensional case where the mood state is simply a scalar. Let \( D(t) \geq 0 \) and \( M(t) \geq 0 \) represent the severity of depression (respectively, mania) at time \( t \geq 0 \). Consider Figure 1a (see caption first), but ignore everything except the red/blue vertical line. Place both \( D(t) \) and \( M(t) \) on this line (one dimension). Thinking of depression and mania as opposites, assume \( D(t) \) is measured on the blue vertical axis, and \( M(t) \) on the red vertical axis. The “×” represents normal/euthymic. If \( D(t) \) starts at normal and its severity increases, mood decreases in the downward direction along the blue line (toward a bottom “pole”). If \( M(t) \) starts at normal and its severity increases, the mood point increases in the upward direction along the red line (toward the other pole). Representing mood in this manner allows for the colloquial manner of discussing “mood swings” going “up and down” that ignores the mixed state.

The “mood state” [1] is represented here via two independent variables at time \( t \geq 0 \), depressive mood \( D(t) \) and manic mood \( M(t) \). Numerical values and scales could result from standard instruments for measuring depression (e.g., BDI or HAM-D; see [3]), mania (e.g., YMRS; see [1]), or the mixed state (e.g., see [26, 42]). Here, we assume that the numeric results from such instruments are aggregated and/or scaled so that they take on values between zero and one. Then, the variables \( D(t) \) and \( M(t) \) can be represented on a standard cartesian plane (two-dimensional plot), with the horizontal being \( D(t) \), and the vertical, \( M(t) \). Then, all combinations of severity levels for depression and mania can be represented
Figure 1: For each figure, we *rotate* the standard two-dimensional axis by $-45^\circ$ to obtain a “mood plane.” (a) One- and two-dimensional representations of mood; (b) The mood state is in a euthymic region (black circle centered at the green dot); (c) Mood trajectory example with mood starting in a manic/mixed state and decreasing to euthymia with mood lability; and (d) Regions on the mood plane of maximal mood variations for each of the bipolar spectrum disorders [1,22].
via $D(t)$ and $M(t)$, with values between zero and one (i.e., $D(t) \in [0, 1]$ and $M(t) \in [0, 1]$) representing, in general, the mixed state when the two variables are nonzero (i.e., $D(t) > 0$ and $M(t) > 0$). In Figure 1, we rotate the standard two-dimensional axis by $-45^\circ$ to obtain a “mood plane.” There are three advantages to this rotation: (i) the resulting mood plane generally corresponds to traditional descriptions of mood disorders, like BD-I, where mania corresponds to “up” and depression corresponds to “down;” (ii) the mood plane highlights not only “bipolarity” of BD-I, but also the mixed state and flat affect poles; and (iii) it conceptualizes euthymia as a type of mixed state where there is a normal mix of emotions that create that state.

Next, consider the mood plane in Figure 1a. Different mood states are represented with black dots. Mood change directions are represented by the arrows (vectors) For example, if the mood state represents low depression, and hypomania (dot, upper-left), the vector represents that depression is staying constant, but mania is increasing. The black dot (again, upper-left plot) in the center/bottom, represents that someone is in a mixed state with more depression than mania, and the vector pointing up means that depression is decreasing at the same rate as mania is increasing. In Figure 1b, the mood state is in a euthymic region (black circle centered at the green dot), with “normal” defined for a person or population (see below). A “mood trajectory” is a time-sequence of mood states. Figure 1c shows a mood trajectory example, with mood starting in a manic/mixed state (upper-right) and decreasing to euthymia, but with mood lability (see, e.g., [43] for a discussion on this case); many other mood trajectories will be considered below.

Figure 1d shows regions on the mood plane of maximal mood variations for each of the bipolar spectrum disorders [1,22]. The general description of disorders on the mood spectrum as vertical lines representing mood variations has, e.g., a longer vertical line for BD-I than, e.g., cyclothymia, as mood swings over a wider range for BD-I (e.g., see pp. 22-23, and Figure 1-1, of [1]). When considering the mood plane, the mood variation region for one disorder (e.g., cyclothymia) is a subset of another disorder (BD-I), but this is only in terms of mood variation. An “ordering” of the illnesses on the spectrum analogous to the one in [1] holds, but now in terms of subsets. The dimensional characterization of mood disorders in Figure 1d is related to the one used in Figure 8.1 on p. 145 of [28].

Let the fixed constants $n_d \in [0, 1]$ and $n_m \in [0, 1]$ represent a point in the mood plane. Referencing this point, other mood features can be added to the mood plane cases in Figure 1:

1. **Euthymia**: For $n_d \in [0, 1]$ and $n_m \in [0, 1]$, we assume that if $D(t) = n_d$ and $M(t) = n_m$ this describes “normal” or “euthymic” mood [1, 3]. The values of $n_d$ and $n_m$ could be specified for an individual via assessment or a population by averaging individual assessments. As an example, in Figure 1 $n_d = n_m = 0.1$ is used to represent the center of a euthymic region;

2. **Extreme mood states**: $D(t) = 1$ or $M(t) = 1$ represents maximally severe depression (respectively, mania), and if $D(t) = M(t) = 1$ this represents a maximally severe mixed state [22];

3. **Mixed states**: Intermediate values of $D(t)$ and $M(t)$ represent mixed mood states. If $n_d = n_m = 0.1$, mixed state examples include: (a) $D(t) = 0.75$ and $M(t) = 0.1$ representing very depressed but no mania as compared to normal (e.g., an MDD state);
(b) \( D(t) = 0.2 \) and \( M(t) = 0.1 \) representing light depression but no mania as compared to normal (e.g., a dysphoric state); (c) \( D(t) = 0.25 \) and \( M(t) = 0.25 \) representing moderate depression as compared to normal and moderate mania as compared to normal (e.g., as in a mixed state in cyclothymia); (d) \( D(t) = 0.1 \) and \( M(t) = 0.4 \) representing no depression as compared to normal but hypomania (e.g., as in BD-II); and (e) \( D(t) = 0.75 \) and \( M(t) = 0.75 \) representing a mixed state (e.g., in BD-I).

4. Absence of depression and/or mania: \( D(t) = 0 \) (\( M(t) = 0 \)) represents the total absence of depression (respectively, mania), and if \( D(t) = M(t) = 0 \) this represents “flat affect” [22]. \( D(t) \geq 0 \) with \( M(t) = 0 \) represents “anhedonia.” \( D(t) = 0 \) with \( M(t) \geq 0 \) represents “hedonia” [1];

2.2 Mood Dynamics and Equilibria

The time units adopted in this work are days. We are considering adults who do not experience full-range mood swings within 24 hours [1]. Even though mood can vary with time, for simplicity we frequently drop the notation for time dependency and simply use \( D \) and \( M \) (similarly, for other variables).

Mood dynamics are represented by the differential equations

\[
\frac{dD}{dt} = S_d(D, M, u_d) \\
\frac{dM}{dt} = S_m(D, M, u_m)
\]

with nonlinear functions \( S_d(D, M, u_d) \) and \( S_m(D, M, u_m) \) specifying the rates of change of mood (derivatives, \( \frac{dD}{dt} \) and \( \frac{dM}{dt} \)), where \( u_d(t) \) and \( u_m(t) \) are the internal/external inputs to the depressive and manic dynamics, respectively, e.g., from stimuli psycho-ophysiological response systems.

Consider the “unforced” mania mood dynamics, that is, without the influence of depressive mood or other external inputs and outputs so that \( \dot{M} = \frac{dM}{dt} = S_m(0, M, 0) \). Define

\[
\dot{M} = b_m \frac{a_m (M - n_m + c_m)^2}{a_m (M - n_m + c_m)^2 + 1} - d_m \frac{f_m (M - n_m - g_m)^2}{f_m (M - n_m - g_m)^2 + 1} - h_m (M - n_m)
\]

Double sigmoid (composed of two sigmoids)

In the depressive case, \( \dot{D} = S_d(D, 0, 0) \) is defined in an analogous manner. Solutions to these differential equations exist, and are unique, since \( S_m \) and \( S_d \) are continuous and satisfy Lipschitz conditions. The double sigmoid concept has been employed to model systems with multiple equilibria, of which possibly the most representative and relevant to our work is the Wilson-Cowan model of the interaction between populations of excitatory and inhibitory neurons [44]. To illustrate why the shape of the nonlinear function \( S_m(0, M, 0) \) represents key features of mood dynamics, consider an example. Let \( n_m = 0.5 \), \( b_m = 0.07 \), \( d_m = b_m \), \( c_m = 0.34 \), \( g_m = c_m \), \( a_m = 29 \), \( f_m = a_m \), and \( h_m = 0.19 \). Figure 2a shows the linear decay line (blue, diagonal, see third term in Equation 2 vs. \( M \)), the sum of the rational functions that compose the double-sigmoid (red, see first two terms in Equation 2) vs. \( M \),
and $\dot{M} = S_m(0, M, 0)$ vs. $M$ (magenta, right-hand-side of Equation 2)). For the linear decay (blue) plot, for a given value of $M \geq n_m$ ($M < n_m$) on the horizontal axis there is a negative (positive) value moving mood down (respectively, up); that is, $-h_m (M - n_m)$ tries to stabilize mood to normal. The first two terms in Equation 2) vs. $M$, the red line, have (i) no influence at $M = n_m$; (ii) an increasing positive (negative) influence on mood change as $M$ moves to intermediate values above (below) $M = n_m$ representing destabilizing effects on mood (e.g., making $\dot{M}$ positive when $M > n_m$ so it increases further); and (iii) a lower positive (negative) influence on mood change as $M$ moves above the peaks in the red line, representing destabilizing effects on mood that are weaker for high values of $\pm M$.

The $\dot{M} = S_m(0, M, 0)$ vs. $M$ case, the magenta line, is the sum of all three right-hand side terms in Equation 2, which are the red line and the negative of the blue line. Notice that by plotting the linear decay vs. $M$ we can see the intersection points in Figure 2a that are the five points on the magenta line that cross zero. These zero points identify $M$ values where $\dot{M} = S_m(0, M, 0) = 0$, that is, where there is no change in mood, up or down (these are “equilibria”). For instance, at $M = n_m$, $\dot{M} = S_m(0, M, 0) = 0$, so that when the person is at normal, and there are no influences from depression or internal/external inputs, then mood will stay at normal.

### 2.3 Basins of Attraction for Mood

To visualize the dynamics in the vicinity of the five equilibria, imagine drawing arrows on the horizontal axis of Figure 2a, with the directions indicating how $M$ will change, as specified by the sign of $\dot{M}$, for each value of $M \in [0, 1]$. For example, for $M$ values just above (below) $n_m = 0.5$, the arrow will point to the left (right) since $\dot{M} = S_m(0, M, 0) < 0$ ($\dot{M} = S_m(0, M, 0) > 0$, respectively) and this shows graphically that $n_m = 0.5$ is an asymptotically stable equilibrium point. Since such arrows will move away from the point to the right of $n_m = 0.5$ where the bottom of the valley exists, it is called an unstable equilibrium point (if the $M$ value is to the left of the bottom of that valley, $M$ will decrease towards $n_m$ but if it is to the right of the bottom of the valley, $M$ will increase, moving away from the valley bottom). Similar analyses works for the other three cases where $S_m(0, M, 0) = 0$, and a full analysis of stability is given below.

For another way to view the dynamics and equilibria consider Equations 1 and 2, and taking a continuous-time gradient optimization perspective on the mania dynamics, we let, for any $M(0) \in [0, 1]$ and all $t \geq 0$,

$$
\frac{dM(t)}{dt} = -\alpha \frac{\partial J(M)}{\partial M} \bigg|_{M=M(t)} = S_m(0, M(t), 0)
$$

where $\alpha > 0$ is a constant “step size” and $J(M)$ is the (“potential”) function to be minimized, one that must be chosen so that the dynamics specified by Equation 1 are matched by this equation. Independent of time $t \geq 0$, for any $M \in [0, 1]$, integrating the two right-hand-side terms of this equation we get

$$
J(M) = \int_0^M \frac{\partial J(\lambda)}{\partial \lambda} d\lambda = -\frac{1}{\alpha} \int_0^M S_m(0, \lambda, 0) d\lambda
$$
(a) Functions that define mania mood change.

(b) Basins of attraction for mania.

Figure 2: (a) Unforced manic mood dynamics functions, and (b) Basins of attraction for equilibria via integration of $S_m(0, M, 0)$ and interpretation of each attractor.
For $\alpha = 1$, the plot of $J(M)$ is shown in Figure 2 where there are “basins of attraction” (valleys) for euthymia, anhedonia, and euphoria (in the depression case, there are basins for euthymia, hedonia, and dysphoria, with some justification for this in [30]). The gradient optimization perspective says that if mood is perturbed from the bottom of one of these basins, it will move to go “down hill” until it reaches the bottom of the basin. Hence, the bottoms of these basis “attract” the mood trajectory $M(t)$ if it is in the vicinity of the bottom of the basin. The peaks on the two hills represent “unstable” points where if $M$ is perturbed even slightly to the left or right, mood will tend to move to the left or right more, and hence $M$ will move away from the peak—the tendency is always to move down the $J(M)$ function if it is not at a peak.

Mood shifts between basins in Figure 2 has been linked to external and internal stimuli such as stressors, sleep and seasonal patterns, and the response to these stimuli by other internal processes like physiological arousal or behavioral activation/inhibition systems. For instance, in Figure 2 if mood $M$ starts at euthymia, it is important to know if other variables (e.g., stress and sleep deficits) can move it out of out of the euthymia basin, to the right, over the hill, then down/farther to the right to end up at euphoria. Alternatively, in the analogous diagram to Figure 2 for depression, if mood $M$ starts at dysphoria, it is important to know if other variables (e.g., an anti-depressant) can influence it to move it out of the dysphoria basin, to the right, over the hill, then down/farther to the right to end up at euthymia. In [45], two main processes are identified in the dynamics of depression: neurobiological processes responsible for mood-congruent cognitive biases in attention, processing, rumination, and self-referential schemes, and attenuated cognitive control to correct these biases. Along those lines, we argue that the size of a basin is mostly affected by the degree of biased processing of relevant stimuli. For example, increased processing of negative stimuli by limbic structures in major depressive disorder, which contributes to a reduced stressor tolerance and a reduced threshold for mood switching, is represented by a narrow normal equilibrium basin and wide basin for dysphoria, in the depressive mood case. Also, a basin’s depth could be correlated to the reinforcing elements of mood disorders that maintains mood in the basin, increasing the duration of abnormal episodes. In this case, biased self-reference schemas, mood congruent attention and maladaptive strategies, such as rumination, could increase the depth of the abnormal equilibria’s basin.

Decreasing the parameters $a_m$ and $f_m$ has a larger effect in increasing the depth ratio between the depth of the abnormal and the normal basins. Therefore, low values of these parameters indicates a higher effect of mood-congruency on the cognitive biases in the abnormal basins. This could increase the duration of episodes of abnormal mood, compared to the duration in normal levels. Decreasing $c_m$ and $g_m$ has a greater effect in displacing the unstable equilibria, decreasing the euthymic basin’s size. Thus, low values of $c_m$ and $g_m$ reflect biased processing and low switching thresholds to abnormal mood. An increase in $b_m$ and $d_m$ results in an increase in both the depth and size of the abnormal basins, thus high values in these parameters represent high severity of the cognitive biases and self-referential schemas that attract mood to the abnormal equilibria, leading to long duration or even chronic episodes due to the difficulty of escaping the basin. Low $h_m$ can be associated with attenuated regulatory processes in the prefrontal cortex that regulate mood [45].
2.4 Mood Trajectories for Various Mood Disorders

The differential equation model is flexible enough to represent several of the most important mood disorders in the DSM-5. In Figure 3, the vector field diagrams for BD-II, cyclothymia, and MDD are presented. By altering the size and depth of the basins of attraction for the equilibria, for both the depressive and manic dimensions (as in Figure 2), we can create equilibria at asymmetric locations in the mood plane. The vector field diagram of BD-II features an equilibrium at a hypomania and major depression, while the cyclothymia vector field depicts equilibria at hypomania, and mild depression, as well as in the mixed state, but with increased basin depths to reflect the extended episodic durations in these disorders. MDD features only two equilibria in the normal and major depression episodes.

3 Nonlinear and Computational Analysis of Mood Dynamics

3.1 Stabilization to Euthymia

Here, we provide a narrative on the meaning of the results in Appendix A, where the main theoretical derivation takes place. Note the assumption in Equation 6 (Appendix A) is equivalent to stating a lower bound on the regulation rate $h_m$, given by

$$h_m > \frac{d_m}{g_m}$$

which is independent of $f_m$. This shows that the maximum slope of the sigmoids in the equilibrium corresponding to euthymia is given by $d_m/g_m$. Hence, this is analogous to saying that when the mood regulation rate is large enough, that it can regulate the cognitive biases, then the multi-stability (e.g., equilibria at euphoria and anhedonia) disappears. Larger $h_m$ can be obtained via psychotherapy that promotes awareness of cognitive biases and tools to regulate emotions and mood, as well as pharmacotherapy that targets biological determinants of the prefrontal cortex activity, like serotonin levels. The meaning and effects of these parameters are discussed more in [46–48].

3.2 Mood Equilibria, Pharmacotherapy, and Triggering Mania

Simulations are provided here with depressive mood and manic mood as state variables. Figure 4a shows the vector field diagram for BD-I, with red arrows indicating the direction of mood change for different points in the mood plane. The blue lines represent trajectories with circles at their initial conditions and squares as their final state at time $T = 10$ days. There are four stable equilibria, namely, normal/euthymia, depression, mania, and mixed states where the trajectories may converge (hedonia and anhedonia are ignored as they produce expected results). Note that the rate of convergence matches the onset times given in the literature, which are, on average, three days for mania and seven days for depression [1].

Figure 4b shows the time trajectories for depressive mood in blue, and manic mood in red when there is a misdiagnosis of major depressive disorder, when it should be BD-I, and
Figure 3: Vector field diagram (grid of red arrows) and trajectories (in blue) for (a) BD-II, (b) Cyclothymia, and (c) MDD.
Figure 4: (a) Vector field diagram and trajectories (in blue) for BD-I, showing 4 basins of attraction at normal, depression, mania and mixed states. (b) Mood episode shift from depression to mania after treatment with antidepressant.

the use of an antidepressant. The initial mood state corresponds to the depression episode equilibrium, and at the second day, an antidepressant is administered for two days, modeled here as a constant negative input in depressive mood and a smaller, positive input in manic mood. Even after stopping the medication, manic and depressive mood are attracted to the equilibrium representing mania, which, according to the DSM-5, constitutes sufficient evidence for a BD-I diagnosis.

4 Conclusions

We introduced a general ODE model for illnesses on the bipolar spectrum. Via Lyapunov stability analysis, we show conditions (related to psychotherapeutic ideas) where someone is guaranteed to return to euthymia if they are perturbed off this mood state. Via computational analysis, we show that there can be a switch from depression to mania when bipolar disorder was misdiagnosed as major depressive disorder, and an anti-depressant is administered without a mood stabilizer.

While in the interest of space it was omitted, it is relatively straightforward to add: (i) an analysis of rapid cycling, (ii) analyses medicine moving the state from one equilibrium to another (or shifting an equilibrium), and (iii) a study of how a mood stabilizer functions for bipolar disorders. The most important, and challenging, direction is to expand the model and validate it against experimental data.
Acknowledgements: We would like to thank Andrew Fu and Jeff Layne for some suggestions.

A Appendix: Technical Results on Stability

Here, we present the main technical result, conditions for global exponential stability of the manic mood equilibrium that corresponds to euthymia:

Theorem 1. Equivalence of qualitative properties Consider the unforced manic mood dynamics given in Equation 2, and assume that the mood profile is symmetric as in Figure 2a, with \( b_m = d_m \), \( f_m = a_m \), and \( g_m = c_m \). Then, \( \dot{M} = n_m \) (euthymia) is an equilibrium point and the qualitative properties of the equilibria of the dynamical system in Equation 2 are equivalent to the qualitative properties of the equilibria of the dynamical system

\[
\dot{M} = -p_{m1} \dot{M}^5 + p_{m2} \dot{M}^3 - p_{m3} \dot{M}
\]

where \( \dot{M} = M - n_m \) and

\[
\begin{align*}
p_{m1} &= f_m^2 h_m \\
p_{m2} &= 2f_m h_m (f_m g_m^2 - 1) \\
p_{m3} &= h_m + f_m g_m (h_m f_m g_m^3 + 2h_m g_m - 4d_m)
\end{align*}
\]

Proof. First, at \( M = n_m \), the rate of change of the unforced manic mood in Equation 2 is

\[
S_m(0, n_m, 0) = b_m a_m c_m^2 - d_m \frac{f_m (-g_m)^2}{f_m (-g_m)^2 + 1}
\]

When \( b_m = d_m \), \( f_m = a_m \), and \( g_m = c_m \), then \( S_m(0, n_m, 0) = 0 \), and \( \dot{M} = n_m \) is an equilibrium point. By making the change of variables \( \dot{M} = M - n_m \) and converting the expression in Equation 2 into a single rational function, we have that

\[
\dot{M} = \frac{-p_{m1} \dot{M}^5 + p_{m2} \dot{M}^3 - p_{m3} \dot{M}}{(f_m (M - g_m)^2 + 1) (f_m (M + g_m)^2 + 1)}
\]

where

\[
\begin{align*}
p_{m1} &= f_m^2 h_m \\
p_{m2} &= 2f_m h_m (f_m g_m^2 - 1) \\
p_{m3} &= h_m + f_m g_m (h_m f_m g_m^3 + 2h_m g_m - 4d_m)
\end{align*}
\]

Note that the denominator of the resulting rational function is always positive and continuous. Thus, the domain of the rational function in Equation 5 is the entire real line. Also, the degree of the numerator of the rational function is greater than the degree of the denominator due to the linear decay term, which implies that there are no horizontal asymptotes in the vector field of \( \dot{M} \). Therefore, the qualitative behavior around the equilibria of the dynamical system in Equation 2, which correspond to the zeros of the rational function, is determined by the qualitative behavior of the dynamical system whose vector field is the numerator of the rational function in Equation 5. \( \square \)
Theorem 2. Stability of Euthymia: Consider the unforced manic mood dynamics given in Equation 2, and assume that the mood profile is symmetric as in Figure 2a, with \( b_m = d_m \), \( f_m = a_m \), and \( g_m = c_m \). Then, the manic mood equilibrium point at \( \bar{M} = n_m \) is globally exponentially stable if

\[
h_m > \frac{d_m}{g_m}, \quad g_m > 0. \tag{6}
\]

Proof. Consider the continuously differentiable, radially unbounded, positive function \( V(\hat{M}) = \frac{1}{2} \hat{M}^2 \) as a Lyapunov function candidate. We employ the qualitative result in Theorem 1, by evaluating the derivative of the Lyapunov function candidate along the trajectories given by Equation 5

\[
\dot{V}(\hat{M}) = \hat{M} \left( -p_{m1} \hat{M}^5 + p_{m2} \hat{M}^3 - p_{m3} \hat{M} \right) = -\hat{M}^2 \left( p_{m1} \hat{M}^4 - p_{m2} \hat{M}^2 + p_{m3} \right)
\]

Our goal is to prove that \( \dot{V}(\hat{M}) < -\alpha \|\hat{M}\|^{\beta} \), for \( \alpha > 0 \) and \( \beta > 0 \), which is equivalent to proving that the roots of the last polynomial expression in the parenthesis are real and positive. To simplify notation, we restate the polynomial in the parentheses as

\[
\alpha = P(\hat{M}) = \hat{M}^4 + q \hat{M}^2 + s \tag{7}
\]

where \( q = -p_{m2}/p_{m1} \), and \( s = p_{m3}/p_{m1} \). In [49] it is shown that the positivity of the polynomial in Equation 7 is guaranteed by showing that: (i) the discriminant of \( P(\hat{M}) \) is strictly positive, and (ii) \( s > q^2/4 \). The assumption in Equation 6 fulfills (ii), thus we concentrate on the discriminant of the quartic polynomial, which is

\[
\Delta_4 = 16q^4s - 128q^2s^2 + 256s^3 = s(4s - q^2)^2
\]

The discriminant is the product of the squares of the differences of the roots of \( P(\hat{M}) \). Note that \( \Delta_4 > 0 \) if \( s > 0 \), which is granted by assumption in Equation 6. Therefore, since both conditions hold, \( \alpha = P(\hat{M}) > 0 \), and the equilibrium \( \bar{M} = n_m \) of the system in Equation 2 is globally exponentially stable.

References


