Artificial Pancreas

1st Ian Snider
Mechanical Engineering Dept.
Washington University in St. Louis
St. Louis, USA
i.snider@wustl.edu

1st Tariq Jassim
Electrical Engineering Dept.
Washington University in St. Louis
St. Louis, USA
j.tariq@wustl.edu

***Abstract*—A glucose-insulin system regulates the blood glucose in the human body to a safe level. This system can become impaired, requiring the need for exogenous insulin delivery. Such an insulin delivery device must be robust to unpredictable disturbance signals. To achieve a robust insulin delivery control system, we design a proportional-integral (PI) controller with state estimation to regulate the blood glucose to 100 mg/dL. A linear approximation of the nonlinear glucose-insulin dynamics is used to find poles for tuning the PI controller. Poles are placed by selecting a time-constant and damping ratio. The stability margins for pole placement are determined from an analysis of the control system's open-loop gain Nyquist plot. A meal disturbance signal is also designed to test the robustness and physiological realism of the control system.**

I. Glucose-Insulin Model

The following represents the simplified glucose-insulin model for a three-dimensional dynamical system,

$$\frac{dG(t)}{dt} = -p_1(G(t) - G_b) - X(t)G(t) + D(t) \quad (1)$$

$$\frac{dX(t)}{dt} = -p_2X(t) + p_3(I(t) - I_b) \quad (2)$$

$$\frac{dI(t)}{dt} = -nI(t) + u(t) \quad (3)$$

Here,

- $G(t)$ is the blood glucose concentration (mg/dL).
- $X(t)$ is the plasma insulin concentration (mU/dL).
- $X(t)$ is the insulin action.
- G_b is the baseline glucose concentration (mg/dL).
- I_b is the baseline plasma insulin concentration (mU/L).
- p_1, p_2, p_3 are the system parameters describing insulin sensitivity and glucose dynamics.
- $u(t)$ is the exogenous insulin delivery rate (mU/min).
- $D(t)$ is the rate of glucose influx from meals (mg/dL/min).

In this analysis, we seek to regulate insulin delivery to a patient by designing a proportional-integral (PI) controller with observer input. The desired value to regulate to is $G_{\text{ref}} = 100$ mg/dL. The reading from a glucose sensor is the only measurement available to the insulin delivery device, thus the output of the model is defined as $y(t) = G(t)$. Additionally, the patient is assumed to have the following parameter values within an uncertainty of 15%,

- p_1 = 0.03 min^{-1}
- p_2 = 0.02 min^{-1}
- p_3 = 0.01 min^{-1}
- n = 0.1 min^{-1}
- G_b = 100 mg/dL
- I_b = 10 mU/L

Alterations to the disturbance function $D(t)$ and variance of the parameters within the uncertainty will provide an analysis for the robustness of the designed PI control law. It is worth noting that we make the assumption $I(0) = 0$, or rather that the patient was in diabetic ketoacidosis immediately prior to being hooked into our system.

II. Dynamics Analysis

A. MATLAB Simulation

Figure 1 contains a MATLAB simulation of the uncontrolled glucose-insulin model time response.

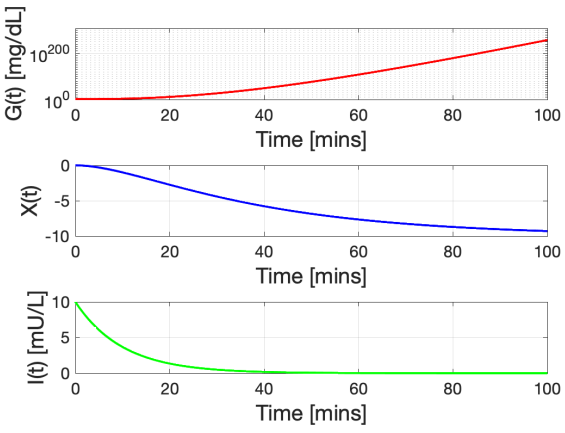


Fig. 1: Uncontrolled time response of $G(t), X(t)$, and $I(t)$. Note that the glucose concentration is simulated on a semi-log plot.

In Fig. 1, the glucose concentration is shown to exponentially increase. This behavior is a product of the simplified mathematical model. While glucose will still release due to the stored sugar in the body, in reality, when not consuming meals, the glucose concentration will not increase this quickly nor reach this level of concentration [1]. As expected, the insulin remains constant since the patient is unable to produce insulin. Also as expected, the negative insulin action stimulates the production of glucose in the body.

B. Characterization of Equilibria

Controller design for the nonlinear system is not straight forward so we instead choose to design a controller based on a Jacobian linearization of the system. The equilibrium point of the undisturbed, uncontrolled glucose-insulin system can be found by determining $G(t), X(t)$, and $I(t)$ where the rate of change is zero. We first define the following state variable,

$$z(t) = \begin{pmatrix} G(t) \\ X(t) \\ I(t) \end{pmatrix} \tag{4}$$

Then, setting the rates of change to zero, we find:

$$\dot{z} = 0 \Rightarrow \quad \begin{aligned} G(t) &= \frac{G_b p_1}{p_1 - \frac{I_b p_3}{p_2}} \\ X(t) &= \frac{-I_b p_3}{p_2} \\ I(t) &= 0 \end{aligned}$$

A closer inspection of this equilibrium point exposes a potential problem. Not only is the glucose-concentration quite far from the baseline value, it is also negative. An evaluation without regard for uncertainty yields $G(t) = -0.604$ mg/dL. The patient may be dissatisfied with a negative glucose concentration. We instead choose to perform the Jacobian linearization around the baseline. The baseline is the point where $G(t) = G_b, X(t) = 0$, and $I(t) = I_b$. Substituting these values into the glucose-insulin model shows that this point is also an equilibrium of the system. Thus, linearizing around this point should provide a more realistic model to design the control law around.

C. Linear Approximation

From the Jacobian of the nonlinear model we find,

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{pmatrix} = \begin{pmatrix} -X - p_1 & -G & 0 \\ 0 & -p_2 & p_3 \\ 0 & 0 & n \end{pmatrix} \tag{5}$$

We define the equilibrium point to be linearized around as,

$$z_{eq} = \begin{pmatrix} G_b \\ 0 \\ I_b \end{pmatrix} \tag{6}$$

Substituting the values of z_{eq} into the Jacobian, we find,

$$A = J(\dot{z})\Big|_{z_{eq}} = \begin{pmatrix} -p_1 & -G_b & 0 \\ 0 & -p_2 & p_3 \\ 0 & 0 & -n \end{pmatrix} \tag{7}$$

$$\tag{8}$$

The linearized model is then complete with,

$$B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \tag{9}$$

To determine the efficacy of designing a control strategy for this model, we examine the controllability and observability. Firstly, for the controllability of the system, the controllability matrix W_C is found to be,

$$\begin{aligned} W_C &= \begin{pmatrix} B & AB & A^2B \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & -G_b p_3 \\ 0 & 0 & -np_3 - p_2 p_3 \\ 1 & -n & n^2 \end{pmatrix} \end{aligned}$$

Aside from edge cases beyond the uncertainty of the provided parameters, W_C is clearly defined by three linearly independent columns thus having $\text{rank}(W_C) = 3$. Thus, the linearized model is controllable. Similarly, the observability matrix W_O is found to be,

$$W_O = \begin{pmatrix} C \\ CA \\ CA^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -p - 1 & -G_b & 0 \\ p_1^2 & G_b p_1 + G_b p_2 & -G_b p_3 \end{pmatrix}$$

Similarly, aside from edge cases beyond the uncertainty of the provided parameters, W_O has $\text{rank}(W_O) = 3$ and the linearized model is thus observable. We may now continue on towards the design of a PI controller with state estimation.

III. PI Controller Design

A. Control Law Definition

The PI control law we seek to implement for the exogenous insulin delivery rate, is defined as,

$$u(t) = -K_p^T \hat{z} - K_i \int_0^t \left(G_{\text{ref}} - \hat{G}(\tau) \right) d\tau \tag{10}$$

where,

$$\hat{z} = \begin{pmatrix} \hat{G} \\ \hat{X} \\ \hat{I} \end{pmatrix} \tag{11}$$

and $\hat{G}(t), \hat{X}(t)$, and $\hat{I}(t)$ are the estimated states for $G(t), X(t)$, and $I(t)$ respectively. Additionally, K_p is the proportional gain and K_i is the integral gain. The proportional control law was chosen as a means to respond to immediate deviations in the blood glucose concentration. The proportional control, while fast, may lead to large steady-state errors. To compensate, integral control was chosen to account for past errors and to mitigate the steady-state error accumulation, thus ensuring that the control law regulates to G_{ref} over time. Additionally, the implementation of an observer allows for estimation of quantities that are hard to measure, such as $X(t)$. Since state estimates are extracted from the nonlinear system, the observer is also able to compensate for oversimplifications due to the linear approximation.

B. Controller & Observer Gains

The approximate LTI model is defined as,

$$\dot{z} = Az + Bu \tag{12}$$

$$y = Cz \tag{13}$$

To solve for the controller gains, an additional state is defined for the integrator term,

$$m(t) = \int_0^t \left(G_{ref} - \hat{G}(\tau) \right) d\tau \tag{14}$$

The state estimation error, e, is defined as,

$$e = z - \hat{z} \tag{15}$$

We then implement a Luenberger observer,

$$\dot{e} = Az + Bu - A\hat{z} - L(y - \hat{y}) \tag{16}$$

$$= (A - LC)e \tag{17}$$

where, L is the observer gain. The dynamics of the system can then be rewritten as,

$$\begin{pmatrix} \dot{z} \\ \dot{m} \\ \dot{e} \end{pmatrix} = \tilde{A} \begin{pmatrix} z \\ m \\ e \end{pmatrix} \tag{18}$$

where,

$$\tilde{A} = \begin{pmatrix} A - BK_p^T & BK_i & 0 \\ -C & 0 & 0 \\ 0 & 0 & (A - BK_p^T + BK_i - LC) \end{pmatrix} \tag{19}$$

The controller gains K_p, K_i, and L, are then found by selecting poles and solving for the controller gains in the characteristic polynomial of $\tilde{A}$ [2]. Tuning pole placement is further described in Sec. III-C.

C. Pole Placement

To determine where exactly we wanted to place our eigenvalues (and thus the poles), we utilized two classic controls principles. Namely, time-constants and damping factors. Recall the time-constant of a LTI system is defined as:

$$T = max_{\lambda_i,\ i \in [0,n]} \frac{-5}{\Re|\lambda_i|} \tag{20}$$

Rearranging this allows us to decide upon a desired time-constant, and determine the real component of the eigenvalues which achieve it:

$$\Re|\lambda_i| = \frac{-5}{T} \tag{21}$$

Using this method allows us to define the time period over which our correction should take effect, but to gain some control over how it gets there (i.e. determine overshooting / undershooting) we need to introduce a damping constant. To justify this, we consider a typical second order dynamical system defined as:

$$\ddot{x} + 2\zeta\omega_n\dot{x} + \omega_n^2 x = 0 \tag{22}$$

This system has the very useful property that both its transient and steady-state behavior can be completely characterized from its eigenvalues, which are given by the quadratic formula derived from the characteristic equation of the system:

$$\lambda_{1,2} = -\zeta\omega_n \pm j\omega_n\sqrt{1 - \zeta^2} \tag{23}$$

where ζ is the damping ratio. From ζ, we can determine the degree of the oscillatory nature of the system by comparing to the following three cases:

1) Under-damped ($0 < \zeta < 1$): The eigenvalues are complex conjugates, and the system exhibits oscillations with a frequency proportional to $\omega_n\sqrt{1 - \zeta^2}$ and an exponential decay governed by $-\zeta\omega - n$.
2) Critically damped ($\zeta = 1$): The eigenvalues are real and equal, resulting in a non-oscillatory system with the fastest return to equilibrium without overshooting.
3) Over-damped ($\zeta > 1$): The eigenvalues are distinct and real, leading to a slower, non-oscillatory return to equilibrium.

By combining these concepts together, we can use our pick of time-constant and damping factor to solve for the natural frequency, ω_n:

$$\Re|\lambda_i| = \frac{-5}{T} = -\zeta\omega_n \Rightarrow \omega_n = \frac{5}{T\zeta} \tag{24}$$

Then we can simply solve for λ as previously defined. However, you may notice there are only two eigenvalues in this approach, yet for our system we need to find seven! Three for our states $G(t), X(t), I(t)$, one for our integral action M and three for our observer. To reduce the complexity of our solution, We decided to define all seven of our eigenvalues relative to these two simple second order ones. Specifically, we defined:

$$\lambda = \begin{bmatrix} \frac{-5}{T} + (\frac{5}{T\zeta}\sqrt{1-\zeta^2})j \\ \frac{-5}{T} - (\frac{5}{T\zeta}\sqrt{1-\zeta^2})j \\ 1.1 \times \lambda_1 \\ 100 \times \lambda_1 \\ -25.5/T \\ -26/T \\ -26.5/T \end{bmatrix} \tag{25}$$

Our motivations for this were intentioned and justified. We matched our first two eigenvalues to that of the second order system, and scaled our third to be slightly larger than the first. This effectively reinforces the rise behavior governed by the original second order system, whose behavior will arise from the placement of $\lambda_1, 2$. Our fourth eigenvalue corresponds to our integral control, and as such we placed it two orders of magnitude away from our proportional poles in the LHP with the intention that it would allow for steady-state error correction without muddying up the transient. Finally, we set our observer poles to be roughly 5 times larger than the real component of our second order system. We need to ensure the observed states converge faster than the autonomous system so our feedback is not overly phase shifted / incorrect (conventional wisdom is 2-10x [2]). The reason we only chose to scale the real component and omit the imaginary component is to hopefully avoid possible resonance / phase interactions between the observer and true state.

Thus we reduced the task of placing seven poles to instead fine tuning 2 parameters, T and ζ. We chose $T = 20$ and $\zeta = 1.1$ for our final controller, for a multitude of reasons. We knew we wanted an over damped system to prevent dangerous spikes and dips that may come with an under-damped system. We also knew that we wanted our time-constant to be smaller than the expected duration of the meal disturbance function so that it would have time to correct before the glucose left the acceptable range. Since our meal disturbance function has a decay rate of $-2t$, we expect meals to continue impacting the glucose for ≈ 90 minutes. Thus we chose a time constant a little less than half of that to account for the non-uniform intensity distribution of the meal. Combining these two factors led us to our choice of T & ζ, where in order to ensure our controller acted appropriately to meals the damping needed to be less than 2 to avoid dangerous spikes and dips. After we determined our eigenvalues we utilized MATLAB's `place()` function to find our actual constants K_p, K_i, L:

$$K_p = \begin{pmatrix} -12.41 \\ 3766 \\ 36.37 \end{pmatrix}$$

$$K_i = 1.4576$$

$$L = \begin{pmatrix} 3.675 \\ -0.0443 \\ -1.622 \end{pmatrix}$$

To test these gains, we first apply our control law to the undisturbed model. Figure 2 contains the system response to the PI controller without disturbances.

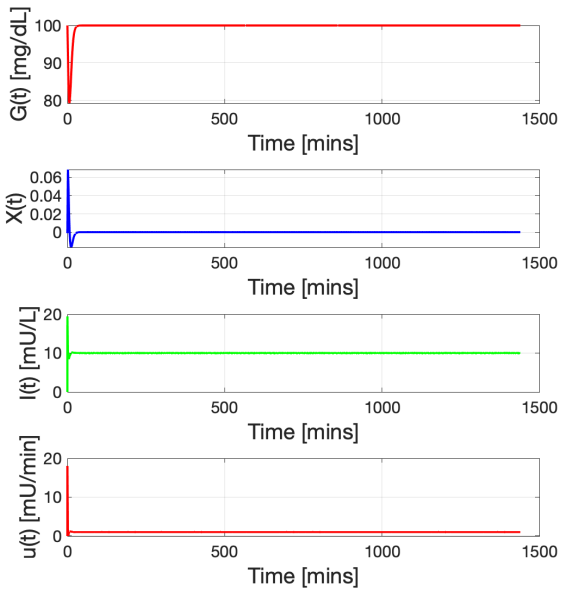


Fig. 2: Time response without meal disturbances of the glucose-insulin dynamics with PI controller and observer.

In contrast with the uncontrolled model, the glucose concentration does not exponentially increase. Additionally, the insulin action is only negative at the start. The integral controller requires time to correctly mitigate the system response, so the somewhat large integral gain produces an initial dip in the glucose concentration.

D. Transfer Functions

Analyzing the linear approximation transfer functions allows for a more tractable evaluation of the loop stability margins providing a characterization of the insulin delivery robustness. The plant transfer function for the LTI system is found as follows,

$$P(s) = \frac{Y(s)}{U(s)} = C(sI - A)^{-1}B \tag{26}$$

$$= \frac{-G_b p_3}{(s + p_1)(s + p_2)(s + n)} \tag{27}$$

Where s is the frequency $j\omega$ and $Y(s)$ and $U(s)$ are the Laplace transforms of the output and control law. The controller transfer function is found as,

$$C(s) = \frac{U(s)}{E(s)} = K_p + \frac{K_i}{s} \tag{28}$$

Where $E(s)$ is the Laplace transform of the error signal. The closed-loop transfer function of the feedback system is then,

$$G_{cl}(s) = \frac{P(s)C(s)}{1 + P(s)C(s)} \tag{29}$$

A pole of the closed-loop system is found when $P(s)C(s) = -1$. The open-loop gain of the system $L(s)$ is then,

$$L(s) = P(s)C(s) \tag{30}$$

Analyzing the loop gain $L(s)$ as a function of the input frequency will allow for a characterization of stability within the linear approximation of the glucose-insulin dynamics.

E. Stability

To evaluate the stability of our system, we turned to the Nyquist stability criterion. By analyzing the Nyquist plot of our open-loop gain $P(s)C(s)$ we can determine if the system is stable, and by how much. To elaborate, we can utilize MATLAB's `margin()` function on our open-loop gain to find both the **Phase Margin** and **Gain Margin** of our model. Recall that the phase-margin is how much phase shift can enter the system before the point $(-1, 0)$ is encircled, and that the gain-margin is the extent to which you can scale your gain (and by extension how far you "slide" your 0 crossings) before $(-1, 0)$ is encircled.

The output of `margin()` is the Bode Plot shown in Fig. 6, in the appendix. We clearly see an infinite Gain margin and roughly 170° of phase margin. By examining the Nyquist Diagram in Fig. 3 we can see that since there are no zero crossings in the LHP, we can scale our gain

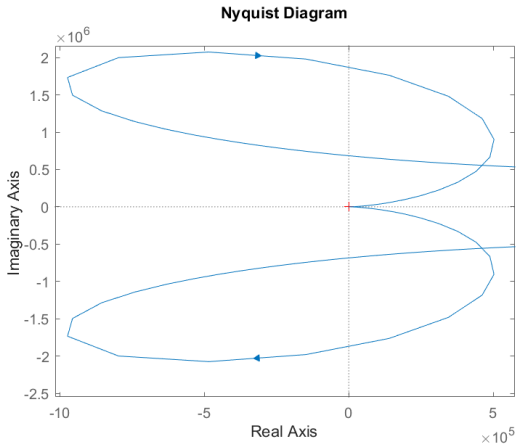


Fig. 3: Nyquist plot of Open-Loop Gain

arbitrarily large without impacting stability. Although, in practical applications this is not fully true, as numerical instability may arise at extremely large values and the robustness of the model will be reduced at larger gains. Though it's enough, at least to justify our 100x scaling for the placement of our integral control's pole. We can also confirm the phase margin, but first take note of the axis scaling of 10^5. If we consider -1 to be at $(0, 0)$, we can envision how large of a rotation must occur before there is a line through it. On this graph the "limiting" line is the "loop-back" on the bottom half plane. If we were to overlay a unit circle onto the image, the straight shot intersection would be around $\frac{15\pi}{8}$. Thus, the phase margin is how long it takes to get from $\frac{15\pi}{8}$ to π, which is $\frac{7\pi}{8}$ or 50°, reasonably close to what MATLAB reports.

Our system should remain stable under most foreseeable operating conditions, as it is robust to both phase and margin shifts. In terms of our actual system robustness, means we can endure a time delay of up to:

$$max(\text{Time Delay}) = \frac{\text{Phase Margin}}{\text{Crossover Frequency}} \tag{31}$$

$$= \frac{\frac{7\pi}{8}}{62}$$

which is ~ 0.0476 minutes or 2.856 seconds.

IV. Disturbance Signal

For the model disturbance signal, we attempt to capture realistic eating habits of a "normal" person. Here, a "normal" person refers to an individual who has a consistent sleep schedule and eats approximately the same number of meals each day at approximately the same time. This signal would not, for instance, capture the eating and sleeping habits of an ESE student during finals week. For the interval from 0 to 12 hours, we randomly distribute 4 meals. For the interval 12 to 16

hours, we randomly distribute a single meal. Then, no meals are consumed for 8 hours.

The magnitude of the glucose influx disturbance signal is a value between 5-20 mg/dL/min. This range is chosen based on the rate of gastric emptying. The typical rate of gastric emptying lies between 1-4 kcal/min which is approximately 0.25-1 g of glucose per minute [1]. For an average human, with approximately 5 L of blood, the gastric emptying corresponds to approximately 5–20 mg/dL/min. Choosing a magnitude of 10 mg/dL/min for each meal, we define,

$$D(t) = \begin{cases} \sum_{i=1}^{4} 10e^{-2(t-a_i)} & t \in [0,12), t \geq a_i \\ 10e^{-2(t-a_5)} & t \in [12,16), t \geq a_5 \\ 0 & t \in [16,24) \end{cases} \quad (32)$$

This model for $D(t)$ represents a time shifted exponential decay for the rate of glucose influx, where a is the randomly distributed peak times within the interval. A quirk of this model is that it may result in superposed disturbances. These superposed disturbances can be thought of as an individual consuming a larger meal that then sustains them for more time than a single normal meal. When examining robustness in Sec. V, we see that this sometimes results in a large glucose concentration. Figure 5, in the appendix, contains a plot with a sample distribution of meal disturbances mitigated by the designed PI control law.

In Fig. 5, meal disturbance signals according to Eq. 32 are mitigated by the PI controller and observer. The PI controller rapidly dispels of the excess glucose. A nominal disturbance of 10 mg/dL/min results in a 20 mg/dL increase of the glucose concentration. The PI controller returns the glucose concentration to G_{ref} = 100 mg/dL in approximately one hour.

V. Robustness to Uncertainty

As stated in Sec. I, the uncertainty of the provided parameters is 15%. To evaluate the sensitivity of our insulin delivery system, simulations with our provided disturbance signal detailed in Sec. IV were run twenty times. Each iteration featured a variant of the parameters within the $\pm 15\%$ tolerance and a unique variation of the disturbance signal. The deviation of each simulation from average parameters was calculated using the following cost function,

$$\begin{aligned} \text{Cost}^i = |p_1 - p_1^i|^2 + |p_2 - p_2^i|^2 + \cdots \\ + |p_3 - p_3^i|^2 + |n - n^i|^2 + \cdots \\ + |G_b - G_b^i|^2 + |I_b - I_b^i|^2 \end{aligned} \quad (33)$$

where the superscript i denotes the parameter value for a given iteration. The square difference of the cost function will be zero if the parameter values have no deviation from the mean parameters specified in Sec. I. Figure 4 contains an analysis of the twenty simulations with the maximum and minimum glucose concentrations measured for a given simulation.

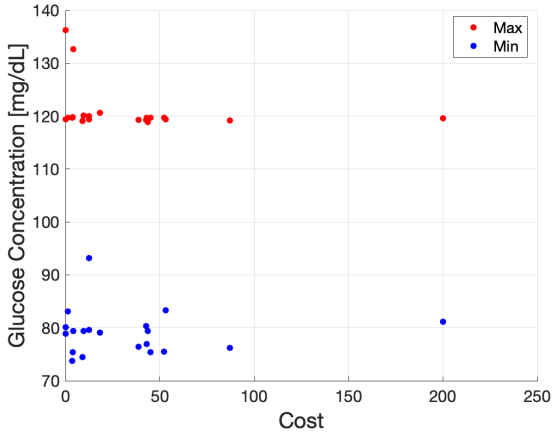


Fig. 4: The simulated maximum and minimum glucose concentrations and their costs.

Figure 4 can be thought of as the meal habits of a certain individual over the course of twenty days. Even with a high cost function output, such as 200, the PI controller maintains maximum and minimum values of 120 mg/dL and 80 mg/dL. These values are within the accepted range. There are however, a few outliers. When the maximum glucose concentration achieved 135 mg/dL an inspection of the corresponding $D(t)$ behavior revealed the superposition of multiple meal disturbances.

VI. Conclusion

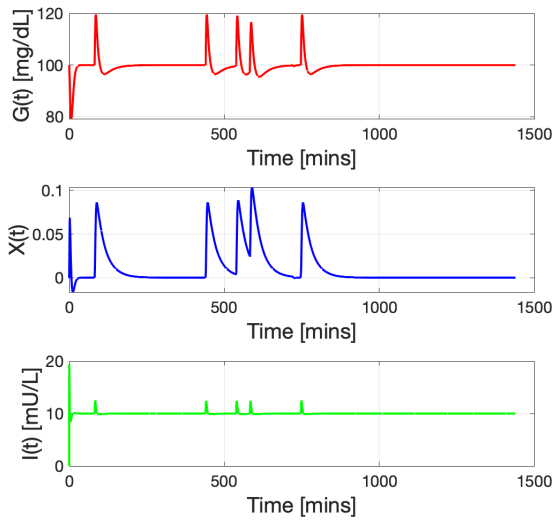
Insulin delivery systems are crucial to the regulation of blood glucose concentration for individuals who are unable to produce their own insulin. The proposed PI control system is robust to almost any magnitude of glucose influx, making it a reasonable solution for commercial use. Our stability analysis suggests a real world system could have a lag of up to 2.5 seconds. However, this control system is computationally expensive. Making it potentially costly to implement on an insulin pump microprocessor. This could also result in a short battery lifetime, becoming an inconvenience to the user. Forgoing the observer would reduce computational expense and may produce a more feasible control law for patient use.

References

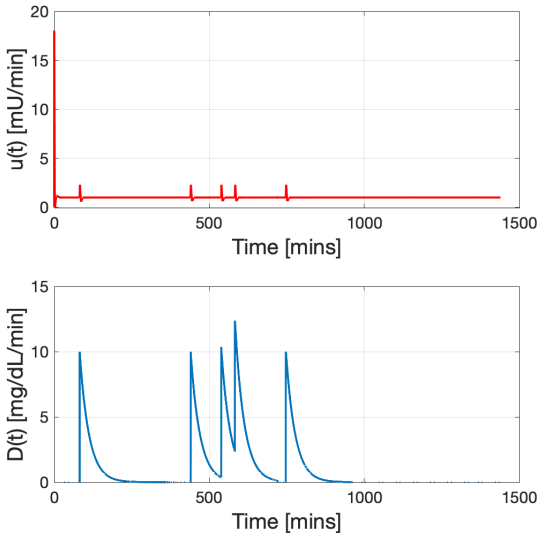
[1] J.J. Holst, F. Gribble, M. Horowitz, C.K. Rayner. Roles of the Gut in Glucose Homeostasis. Diabetes Care, 2016. doi: 10.2337/dc16-0351. PMID: 27222546

[2] K.J. Astrom and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers*, 2nd ed. Princeton University Press, 2021

VII. Appendix



(a) Controlled time response of $G(t)$, $X(t)$, and $I(t)$. Plotted over the course of twenty-four hours (1440 mins).



(b) Insulin delivery rate (top) and meal disturbance signal (bottom).

Fig. 5: Time response of the glucose-insulin dynamics with PI controller and observer.

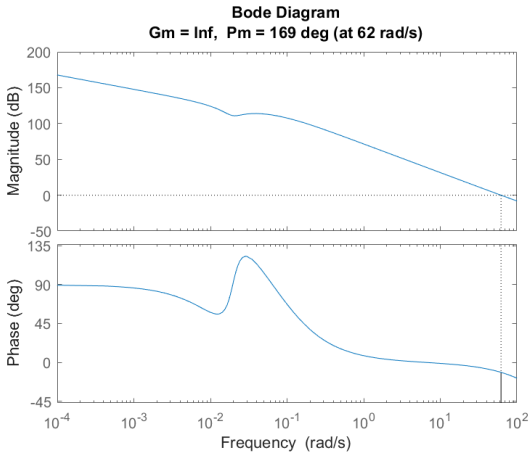


Fig. 6: Output of MATLAB's *margin()* on Open-Loop Gain