Working Notes on Glucose–Insulin Modeling

Based on the Dalla Man (2007) Meal Simulation Model

Krishna

January 8, 2026

Contents

1 Reduced-Order Glucose–Insulin Model

1.1 State Variables

We consider a reduced-order representation of human glucose–insulin dynamics with three state variables:

$$\begin{aligned} G(t) &\quad \text{Plasma glucose concentration (mg/dL),} \\ I(t) &\quad \text{Plasma insulin concentration (mU/L),} \\ X(t) &\quad \text{Remote insulin action (min}^{-1}\text{).} \end{aligned}$$

The model is formulated directly in terms of measurable concentrations, avoiding internal mass compartments that are not observable in typical experimental settings.

1.2 Glucose Dynamics

Plasma glucose dynamics are governed by

$$\dot{G}(t) = -(S_G + k_{cl})\big(G(t) - G_b\big) - X(t)\big(G(t) - G_b\big) + \mathrm{Ra}(t), \tag{1}$$

where:

- S_G denotes glucose effectiveness,
- k_{cl} denotes non–insulin–dependent glucose clearance,
- G_b is the basal glucose concentration,
- $X(t)$ represents delayed insulin action,
- $\mathrm{Ra}(t)$ is the rate of glucose appearance from the gastrointestinal tract.

The glucose equation is expressed relative to basal concentration to ensure an analytically consistent fasting equilibrium and to prevent spurious hypoglycemic drift.

1.3 Insulin Dynamics

Plasma insulin dynamics are modeled as

$$\dot{I}(t) = -k_I\big(I(t) - I_b\big) + \beta\,\max\big(G(t) - \theta,\, 0\big) + u_I(t), \tag{2}$$

where:

- k_I is the insulin clearance rate,
- I_b is the basal insulin concentration,
- β is the glucose–stimulated insulin secretion gain,
- θ is the glucose secretion threshold,
- $u_I(t)$ represents exogenous insulin administration.

Insulin secretion is activated only when glucose exceeds a threshold, reflecting pancreatic β–cell response while preserving model simplicity and identifiability.

1.4 Insulin Action Dynamics

Delayed insulin action is represented by a single remote compartment:

$$\dot{X}(t) = -k_X X(t) + \alpha\big(I(t) - I_b\big), \tag{3}$$

where:

- k_X is the insulin action decay rate,
- α is the insulin–to–action gain.

This formulation captures the dominant delay between plasma insulin and its effect on glucose utilization while avoiding over-parameterization.

1.5 Gastric Glucose Appearance Operator

Glucose appearance from a meal is modeled using a mass-conserving operator rather than additional dynamic states:

$$\mathrm{Ra}(t) = D \cdot \mathrm{GammaPDF}\big(t;\, k_g, \theta_g\big), \tag{4}$$

where D is the ingested glucose mass and $\mathrm{GammaPDF}(\cdot)$ denotes the Gamma probability density function with fixed shape k_g and scale θ_g.

By construction,

$$\int_0^\infty \mathrm{Ra}(t)\, dt = D, \tag{5}$$

ensuring exact mass conservation without introducing additional unobservable states.

1.6 Basal Equilibrium

The model admits an analytically guaranteed fasting equilibrium at

$$(G, I, X) = (G_b, I_b, 0), \tag{6}$$

which holds identically in the absence of meal intake and exogenous insulin. This equilibrium is enforced structurally rather than through numerical calibration.

1.7 Parameters and Identifiability

The parameter set is given by

$$\{S_G,\, k_I,\, k_X,\, \alpha,\, \beta,\, \theta\},$$

while k_{cl} is fixed to a physiologically plausible value due to its structural non-identifiability with S_G in Eq. (1). Structural and practical identifiability of the remaining parameters is verified through sensitivity analysis, Fisher Information Matrix evaluation, and profile likelihood analysis.

2 Phase 2: Full-Model Identifiability Analysis (Foundational Phase)

2.1 Motivation and Logical Priority

Before addressing questions such as which states to retain, how to reduce the model, or how to design a reduced-order representation, a more fundamental question must be resolved:

> **Is the full Dalla Man model identifiable from plasma glucose measurements in principle?**

This question is logically prior to all subsequent modeling steps, including:

- reduced-order modeling (ROM),
- state elimination or aggregation,
- time-scale separation arguments,
- numerical parameter estimation.

If identifiability fails at this stage, no amount of clever reduction, reparameterization, or numerical optimization can recover the original full-model parameterization. Phase 2 therefore constitutes a *structural stress test of the model itself*, rather than a test of data quality, experimental design, or estimation algorithms.

2.2 Core Constraints Imposed

Throughout Phase 2, several constraints were deliberately fixed in order to isolate intrinsic limitations of the model:

- **Output**: plasma glucose $G_p(t)$ only.
- **Model structure**: the full, unmodified Dalla Man equations.
- **Operating regime**: normoglycemic, postprandial conditions.
- **Assumptions**: optimistic, best-case assumptions whenever possible.

Under these constraints, any failure of identifiability must be attributed to the algebraic and dynamical structure of the model, not to limitations of the data.

2.3 Phase 2.1: Sensitivity Analysis

2.3.1 Question Addressed

The sensitivity analysis addresses the following question:

> *Does plasma glucose respond independently to individual model parameters?*

If plasma glucose exhibits negligible or redundant response to variations in a parameter, that parameter is effectively invisible to glucose-based measurements.

2.3.2 Methodology

Normalized local sensitivities of plasma glucose with respect to each parameter p_j were computed as

$$S_j(t) = \frac{p_j}{G(t)} \frac{\partial G(t)}{\partial p_j}. \tag{7}$$

Sensitivities were evaluated under a variety of conditions:

- fasting conditions,
- postprandial excitation,
- multiple meal sizes,
- different observation time windows.

The sensitivities were then integrated over time to obtain global influence measures for each parameter.

2.3.3 Observations

The sensitivity analysis revealed several consistent patterns:

1. **Extreme dominance** of a small subset of parameters, notably V_G, k_{p1}, k_1, and k_2.
2. Many parameters exhibited **near-zero sensitivity** or were active only transiently.
3. Strong **correlation clusters** were observed, with gastric parameters moving together and insulin-related parameters compensating one another.

2.3.4 Key Inference

The sensitivity directions were found to be highly non-independent. Multiple parameters produced nearly identical glucose responses, indicating that plasma glucose does not encode parameter-specific information in a separable manner. While sensitivity analysis alone cannot prove non-identifiability, it provides the first strong indication of severe parameter redundancy.

2.3.5 Implication

Even before considering noise, estimation error, or numerical algorithms, plasma glucose does not "see" individual parameters independently. This motivates the use of information-theoretic diagnostics.

2.4 Phase 2.2: Fisher Information Matrix Analysis

2.4.1 Question Addressed

The Fisher Information Matrix (FIM) analysis addresses the question:

How many independent parameter combinations can glucose data actually identify?

This constitutes a geometric extension of the sensitivity analysis.

2.4.2 Methodology

A sensitivity matrix was constructed as

$$S(t) = \left[\frac{\partial G}{\partial p_1}, \ldots, \frac{\partial G}{\partial p_m} \right]. \tag{8}$$

The Fisher Information Matrix was then computed as

$$\mathcal{I} = \int_0^T S(t)^\top S(t), dt. \tag{9}$$

Eigenvalues, condition numbers, and dominant information directions were analyzed. Deliberately rich excitation and noise-free assumptions were employed to construct a best-case identifiability scenario.

2.4.3 Observations

The Fisher Information Matrix exhibited the following properties:

1. Only one or two dominant eigenvalues.
2. Remaining eigenvalues collapsed by several orders of magnitude.
3. Severe ill-conditioning of the matrix.
4. Increasing excitation amplified dominant eigenvalues while further suppressing the smallest ones.

2.4.4 Key Inference

Plasma glucose carries information in a very low-dimensional subspace of the parameter space. Most parameters lie along near-null directions that cannot be recovered, regardless of experimental richness.

2.4.5 Critical Insight

If increasing excitation worsens identifiability, the limitation is structural rather than experimental.

2.5 Phase 2.3: Profile Likelihood Analysis

2.5.1 Question Addressed

Profile likelihood analysis addresses the question:

Can parameters be uniquely estimated in practice, even if sensitivities exist?

This test bridges theoretical identifiability and numerical estimation.

2.5.2 Methodology

Synthetic glucose data were generated from the full model and corrupted with 10

2.5.3 Observations

The resulting profile likelihoods exhibited:

1. Flat or weakly curved profiles.
2. Unbounded and asymmetric confidence intervals.
3. Profiles that failed to cross statistical confidence thresholds.

2.5.4 Key Inference

Even under optimistic conditions, multiple parameter sets produce indistinguishable glucose trajectories. Parameters are therefore practically non-identifiable.

2.5.5 Significance

Profile likelihood analysis rules out numerical artifacts as the source of identifiability failure.

2.6 Phase 2.4: Structural Identifiability Analysis (GenSSI)

2.6.1 Question Addressed

Structural identifiability analysis addresses the strongest question:

> *Is non-identifiability inherent in the equations themselves?*

2.6.2 Methodology

The GenSSI framework was employed to analyze generating series based on Lie derivatives of the output. Identifiability tableaux and rank growth with increasing derivative order were examined.

2.6.3 Observations

The analysis revealed:

1. Parameters appearing only through composite expressions (e.g., G_p/V_G, $k_{abs}Q_{gut}$).
2. Linear dependence among identifiability tableau columns.
3. Slow rank growth with increasing derivative order.
4. No parameter becoming uniquely identifiable.

2.6.4 Key Inference

Non-identifiability is structural. No experiment, estimator, or data volume can resolve it. Plasma glucose collapses multiple physiological mechanisms into identical algebraic effects.

2.7 Global Synthesis of Phase 2

Across sensitivity analysis, Fisher information analysis, profile likelihoods, and structural identifiability testing, the conclusions are consistent:

- Parameters are intrinsically coupled.
- Information content is fundamentally low-dimensional.
- Compensation mechanisms are structural.
- Plasma glucose alone is an insufficient output for the full model.

Consequently, the full Dalla Man model is:

- **structurally non-identifiable**,
- **practically non-identifiable**,
- **over-parameterized for glucose-only data**.

This conclusion is overdetermined and does not rely on any single diagnostic.

2.8 Role of Phase 2 in the Overall Workflow

Phase 2 establishes the information-theoretic impossibility of recovering the full model from glucose measurements alone. As a result:

- Reduced-order modeling becomes necessary rather than optional.
- State elimination is justified rather than heuristic.
- Parameter lumping is forced rather than arbitrary.
- Subsequent dynamical diagnostics acquire clear meaning.

> **Reduced-order modeling is a response to an information-theoretic impossibility, not a modeling preference.**

3 Phase 4: Empirical Dynamical Diagnostics on the Full Model

3.1 Motivation and Logical Role

Phase 4 is entered only after the conclusions of Phase 2 have been firmly established. Phase 2 demonstrated that the full Dalla Man model is structurally and practically non-identifiable from plasma glucose measurements alone. As a result, parameter recovery is no longer a meaningful objective.

The purpose of Phase 4 is therefore fundamentally different from that of Phase 2. Rather than asking whether parameters can be identified, Phase 4 asks the following conditional question:

> **Given that the full model cannot be identified, which internal dynamical components actually matter for shaping plasma glucose dynamics?**

This phase is explicitly *state-centric* rather than parameter-centric. The goal is to determine which states participate meaningfully in the transmission of physiological inputs to the glucose output, which states evolve on independent time scales, and which states are dynamically slaved, redundant, or shielded from observation. These questions are essential for constructing a reduced-order model that is both defensible and faithful to the observed glucose dynamics.

3.2 Conceptual Shift from Phase 2

The transition from Phase 2 to Phase 4 represents a deliberate change in perspective:

- Phase 2 focuses on algebraic structure and identifiability limits.
- Phase 4 focuses on realized nonlinear dynamics along physiologically relevant trajectories.
- Phase 2 is concerned with parameters.
- Phase 4 is concerned with states and information flow.

This distinction is crucial. A state may be structurally present and physiologically meaningful, yet dynamically irrelevant to plasma glucose under the operating regimes of interest. Phase 4 is designed to expose this distinction quantitatively.

3.3 Phase 4.1: Test I — Empirical Input–Output Energy Ranking

3.3.1 Question Addressed

The first dynamical diagnostic addresses the question:

> *Which internal states actively transmit physiologically admissible inputs to the plasma glucose output?*

A state is considered dynamically relevant for reduced-order modeling if it is both excited by inputs and if that excitation is reflected in the measured output. States that fail either criterion are dynamically redundant from the perspective of glucose-only modeling.

3.3.2 Methodology

Two physiologically meaningful input experiments were designed:

1. an insulin perturbation experiment, intended to excite insulin kinetics and insulin action pathways,
2. a meal-related perturbation experiment, intended to excite gastric emptying and intestinal absorption dynamics.

For each state $x_i(t)$, the following empirical quantities were computed:

$$W_{c,i} = \int_0^T x_i(t)^2, dt, \; W_{o,i} \qquad = \int_0^T \big(x_i(t), G_p(t)\big)^2, dt. \tag{10}$$

These represent empirical controllability and observability energies, respectively. A joint input–output influence metric was then defined as

$$I_i = \sqrt{W_{c,i} W_{o,i}}, \tag{11}$$

and normalized across all states to obtain relative influence rankings.

3.3.3 Observations

The resulting influence distribution was extremely sparse. Only the gastrointestinal states exhibited non-negligible influence:

- Q_{sto2} (triturated stomach compartment) dominated overwhelmingly,
- Q_{sto1} (solid stomach compartment) contributed at a substantially lower level,
- Q_{gut} (intestinal glucose compartment) contributed marginally.

All insulin-related states (I_p, I_l, I_1, I_d, X, Y) exhibited influence values effectively indistinguishable from zero. Even the glucose states G_p and G_t were weak relative to the gastric states.

3.3.4 Inference

Plasma glucose dynamics are overwhelmingly dominated by the gastrointestinal subsystem under the tested operating regime. Large portions of the insulin subsystem are dynamically slaved and do not independently transmit input–output energy. This provides the first strong state-level justification for aggressive reduction.

3.4 Phase 4.2: Test II — Time-Scale Separation Analysis

3.4.1 Question Addressed

The second dynamical diagnostic asks:

> *Do subsets of states evolve on widely separated time scales, indicating fast slaving or quasi-steady behavior?*

3.4.2 Methodology

A baseline postprandial trajectory was generated using the full nonlinear model. At multiple time points along this trajectory, the system was locally linearized by computing the Jacobian matrix $J(x(t))$. The eigenvalues of the Jacobian were converted to local linear time constants via

$$\tau_i = -\frac{1}{\Re(\lambda_i)}. \tag{12}$$

This procedure was repeated at early postprandial, peak glucose, and late recovery phases to assess persistence of time-scale separation.

3.4.3 Observations

Across all operating points, the fastest time constants were on the order of one minute, while the slowest time constants were on the order of one to two hours. The resulting separation ratio consistently exceeded 10^2. Importantly, this separation was observed throughout the trajectory and was not confined to a single transient phase.

3.4.4 Inference

The model exhibits strong and persistent fast–slow structure. Many states rapidly equilibrate and subsequently track slow manifolds determined by a small subset of dominant dynamics. This provides dynamic justification for quasi-steady approximations and state aggregation.

3.5 Phase 4.3: Test III — Excitation-Rich Dynamical Stress Test

3.5.1 Question Addressed

The third diagnostic asks:

> *Does increasing excitation reveal additional independent dynamical degrees of freedom relevant to glucose?*

3.5.2 Methodology

Deliberately rich insulin and meal perturbations were applied to the system, producing glucose trajectories with increased temporal complexity. Sensitivities and Fisher information structures were re-examined under these conditions.

3.5.3 Observations

While dominant dynamical modes were amplified under richer excitation, weak modes became even less visible. No new independent directions emerged, and the effective dynamical dimensionality remained unchanged.

3.5.4 Inference

Additional excitation does not activate hidden glucose-relevant dynamics. Instead, it reinforces the dominance of already-active pathways. This rules out insufficient excitation as an explanation for state redundancy.

3.6 Phase 4.4: Test IV — Output Map Sensitivity Analysis

3.6.1 Question Addressed

The final dynamical diagnostic asks:

> *Which states directly shape the instantaneous dynamics of plasma glucose?*

3.6.2 Methodology

Along the baseline trajectory, the sensitivity of the glucose derivative with respect to each state was computed as

$$S_i(t) = \frac{\partial \dot{G}_p(t)}{\partial x_i}. \tag{13}$$

Mean absolute sensitivities were then aggregated over time to produce relative influence rankings.

3.6.3 Observations

The strongest sensitivities corresponded to the glucose states G_p and G_t. Moderate indirect influence was observed from Q_{gut}. Insulin-related and auxiliary states exhibited negligible influence on $\dot{G}_p$.

3.6.4 Inference

Many states affect glucose only through collapsed pathways and do not appear explicitly in the output map. Even dynamically active states may therefore be output-invisible.

3.7 Global Synthesis of Phase 4

Across all Phase 4 diagnostics, a consistent picture emerges:

- Only a small subset of states carry significant input–output energy.
- Strong fast–slow separation leads to state slaving.
- Insulin dynamics are largely indirect and collapsed from the glucose perspective.
- The effective dynamical dimension of the system is extremely low.

3.8 Role of Phase 4 in the Overall Workflow

Phase 4 establishes that large portions of the full Dalla Man state space are dynamically redundant with respect to plasma glucose. Combined with the identifiability results of Phase 2, this phase provides the final justification for constructing a reduced-order model that preserves only the dominant glucose-shaping dynamics.

> **Phase 2 proves that the full model cannot be identified; Phase 4 proves that much of it does not dynamically matter.**

4 Profile Likelihood: Complete Mathematical Construction with a Minimal Example

4.1 Problem Statement

The goal of profile likelihood analysis is to answer the following precise question:

> If one parameter is fixed to a chosen value, how well can the model fit the observed data after all remaining parameters are optimally adjusted?

This question is strictly stronger than both sensitivity analysis and best-fit parameter estimation. Sensitivity analysis measures local responsiveness, while best-fit estimation identifies a single optimal parameter vector. In contrast, profile likelihood directly probes the *information content of the data* by testing whether deviations in one parameter can be compensated by others without significantly degrading the quality of fit.

4.2 Definition of Profile Likelihood

Profile likelihood is a *function of a single parameter*. For a chosen parameter, it returns the minimum achievable misfit after re-optimizing all remaining parameters. In other words, it maps each fixed value of the parameter of interest to the best possible fit that the model can produce.

Thus, profile likelihood is not a probability distribution, but a deterministic function:

$$\text{parameter value} \longrightarrow \text{minimum achievable } \chi^2.$$

4.3 Minimal Model Demonstrating Parameter Compensation

To illustrate the construction transparently, we consider the simplest possible model that exhibits parameter compensation.

Model

$$y = a + b$$

- Parameters: a, b
- Model output: y

4.4 Observed Data

We consider a single measurement:

$$y_{\text{meas}} = 10,$$

with measurement noise standard deviation

$$\sigma = 1.$$

The use of a single data point is intentional, as it exposes identifiability issues in their most elementary form.

4.5 Construction of the Chi-Square Objective Function

The model prediction is $y_{\text{sim}} = a + b$. The residual (error) is therefore

$$e(a, b) = y_{\text{meas}} - (a + b).$$

Under Gaussian noise assumptions, the chi-square objective function is

$$\chi^2(a, b) = \left(\frac{y_{\text{meas}} - (a + b)}{\sigma} \right)^2 = (10 - a - b)^2.$$

This function defines a surface in the two-dimensional parameter space (a, b).

4.6 Global Minimum of the Chi-Square Function

The chi-square function is minimized when

$$a + b = 10.$$

At these points,

$$\chi^2_{\text{min}} = 0.$$

There are infinitely many parameter combinations satisfying this condition, including:

$$(a, b) = (2, 8), \quad (5, 5), \quad (9, 1).$$

This immediately suggests non-identifiability, which we now formalize using profile likelihood.

4.7 Profile Likelihood Construction for Parameter a

Step A: Fix the Parameter of Interest

We fix parameter a to a constant value:

$$a = \alpha,$$

where α is treated as a prescribed number.

Step B: Re-optimize Remaining Parameters

With a fixed, the remaining free parameter is b. The profile likelihood is defined as the minimum chi-square achievable over b:

$$\chi^2_{\text{PL}}(\alpha) = \min_b (10 - \alpha - b)^2.$$

Step C: Explicit Minimization

Differentiating with respect to b:

$$\frac{d}{db}(10 - \alpha - b)^2 = -2(10 - \alpha - b).$$

Setting the derivative equal to zero yields

$$10 - \alpha - b = 0,$$

so that

$$b^{(\alpha)=10-\alpha.}$$

Step D: Evaluate the Profile Likelihood

Substituting $b^{(\alpha)}$ into the chi-square function gives

$$\chi^2_{\mathrm{PL}}(\alpha) = (10 - \alpha - (10 - \alpha))^2 = 0.$$

4.8 Final Profile Likelihood Expression

The profile likelihood for parameter a is therefore

$$\boxed{\chi^2_{\mathrm{PL}}(a) = 0 \quad \forall a.}$$

The corresponding relative chi-square is

$$\Delta\chi^2_{\mathrm{PL}}(a) = \chi^2_{\mathrm{PL}}(a) - \chi^2_{\mathrm{min}} = 0.$$

4.9 Interpretation

This result implies that, regardless of the value to which parameter a is fixed, the model can always adjust parameter b to achieve a perfect fit to the data. Consequently:

- The chi-square value never increases,
- The relative chi-square $\Delta\chi^2$ never crosses the 95% confidence threshold of 3.84,
- No finite confidence interval for a exists.

We therefore conclude:

$$\boxed{\text{Parameter } a \text{ is practically non-identifiable.}}$$

This behavior is directly analogous to the non-identifiability observed for the insulin action delay parameter k_i in the Dalla Man model, albeit in a higher-dimensional and nonlinear setting.

4.10 Contrast Case: Partial Identifiability

If the model were instead

$$y = a + b^2,$$

then fixing a would not always permit perfect compensation by adjusting b. In this case, the chi-square would increase beyond the noise tolerance on one or both sides of the optimum, leading to a bounded confidence interval. This behavior is analogous to the partial identifiability observed for the parameter k_{abs}.

4.11 General Definition of Profile Likelihood

For a general parameter vector $\theta = (\theta_i, \theta_{-i})$, the profile likelihood is defined as

$$\boxed{\chi^2_{\mathrm{PL}}(\theta_i) = \min_{\theta_{j\neq i}} \chi^2(\theta).}$$

Confidence intervals are obtained by enforcing the likelihood-ratio criterion:

$$\boxed{\Delta\chi^2_{\mathrm{PL}}(\theta_i) \leq 3.84 \quad (95\% \text{ confidence, one parameter}).}$$

4.12 Big-Picture Takeaway

Profile likelihood is not a plotting heuristic, a numerical trick, or an optimizer artifact. It is a controlled numerical experiment that tests the model's ability to compensate forced parameter changes. For this reason, it is widely regarded as the gold standard for practical identifiability analysis.

5 Structural Identifiability: A Complete Worked Example

This section presents a complete, self-contained demonstration of the concept of *structural identifiability* using the simplest possible dynamical system. All definitions, calculations, and conclusions are derived explicitly, without assuming prior knowledge.

5.1 Problem Statement

Consider a dynamical model with unknown parameters. Assume that the system output is observed perfectly, with infinite temporal resolution and no measurement noise.

The fundamental question of *structural identifiability* is:

> From perfect output data alone, can the true parameter values be uniquely determined *in principle*?

If the answer is negative, the model is said to be *structurally non-identifiable*. In such a case, no experiment design, optimization method, or increase in data quality can resolve the ambiguity.

5.2 Definition of Structural Identifiability

Let

- p denote the unknown parameter(s),
- $x(t)$ denote the system state,
- $y(t)$ denote the measured output.

The model is said to be *structurally identifiable* if and only if

$$y(t; p_1) \equiv y(t; p_2) \quad \forall t \quad \Longrightarrow \quad p_1 = p_2. \tag{14}$$

If two distinct parameter values generate identical output trajectories for all time, they are *indistinguishable*, and the model is structurally non-identifiable.

5.3 Minimal Dynamical Model

We consider the simplest continuous-time system capable of exhibiting identifiability phenomena:

$$\dot{x}(t) = -k\, x(t), \tag{15}$$

$$y(t) = x(t), \tag{16}$$

where

- $x(t)$ is the (hidden) state,
- k is an unknown constant parameter,
- $y(t)$ is the measured output.

The initial condition is unknown:

$$x(0) = x_0. \tag{17}$$

The vector of unknown quantities is therefore

$$\theta = \{k,\ x_0\}. \tag{18}$$

5.4 Explicit Solution and Input–Output Map

The differential equation admits the closed-form solution

$$x(t) = x_0 e^{-kt}. \tag{19}$$

Since $y(t) = x(t)$, the input–output map is

$$y(t) = x_0 e^{-kt}. \tag{20}$$

This expression contains all information available from perfect observation of the output.

5.5 Structural Identifiability Test

Assume that two distinct parameter sets

$$(k_1, x_{0,1}) \neq (k_2, x_{0,2}) \tag{21}$$

produce identical outputs:

$$x_{0,1} e^{-k_1 t} = x_{0,2} e^{-k_2 t} \quad \forall t. \tag{22}$$

Dividing both sides yields

$$\frac{x_{0,1}}{x_{0,2}} = e^{(k_1 - k_2)t}. \tag{23}$$

This equality can hold for all t if and only if

$$k_1 = k_2 \quad \text{and} \quad x_{0,1} = x_{0,2}. \tag{24}$$

Conclusion. Both the parameter k and the initial condition x_0 are *structurally identifiable*. This constitutes a positive identifiability result.

5.6 Equivalent Derivation via Output Derivatives (GenSSI Approach)

The same conclusion can be obtained without solving the differential equation explicitly, using output derivatives evaluated at $t = 0$.

5.6.1 Output Derivatives

The output and its derivatives are

$$y(0) = x_0, \tag{25}$$
$$\dot{y}(t) = \dot{x}(t) = -kx(t) \quad \Rightarrow \quad \dot{y}(0) = -kx_0, \tag{26}$$
$$\ddot{y}(t) = -k\dot{x}(t) = k^2x(t) \quad \Rightarrow \quad \ddot{y}(0) = k^2x_0. \tag{27}$$

5.6.2 Identifiability Coefficients

Collecting the coefficients yields

$$\Phi = \begin{bmatrix} x_0 \\ -kx_0 \\ k^2x_0 \end{bmatrix}. \tag{28}$$

5.6.3 Identifiability Jacobian

Define the unknown vector

$$\theta = \begin{bmatrix} x_0 \\ k \end{bmatrix}. \tag{29}$$

The identifiability Jacobian is

$$J = \frac{\partial \Phi}{\partial \theta} = \begin{bmatrix} 1 & 0 \\ -k & -x_0 \\ k^2 & 2kx_0 \end{bmatrix}. \tag{30}$$

For generic values $k \neq 0$ and $x_0 \neq 0$, this matrix has full rank:

$$\operatorname{rank}(J) = 2. \tag{31}$$

Conclusion. Since the rank equals the number of unknowns, the model is structurally identifiable.

5.7 A Structurally Non-Identifiable Variant

Now modify only the output equation:

$$\dot{x}(t) = -k\,x(t), \tag{32}$$
$$y(t) = c\,x(t), \tag{33}$$

where c is an additional unknown parameter.

The unknown vector is now

$$\theta = \{k,\ c,\ x_0\}. \tag{34}$$

The output becomes

$$y(t) = cx_0e^{-kt}. \tag{35}$$

5.7.1 Key Observation

Only the product cx_0 appears. Define

$$A = cx_0. \tag{36}$$

Then

$$y(t) = Ae^{-kt}. \tag{37}$$

Different parameter combinations such as

$$(c, x_0) = (2, 5), \tag{38}$$

$$(c, x_0) = (5, 2), \tag{39}$$

produce the same value $A = 10$ and therefore identical outputs for all time.

5.8 Jacobian Proof of Non-Identifiability

The identifiability coefficients are

$$\Phi = \begin{bmatrix} cx_0 \\ -kcx_0 \\ k^2cx_0 \end{bmatrix}. \tag{40}$$

The Jacobian columns corresponding to c and x_0 are

$$\frac{\partial \Phi}{\partial c} = x_0 \begin{bmatrix} 1 \\ -k \\ k^2 \end{bmatrix}, \tag{41}$$

$$\frac{\partial \Phi}{\partial x_0} = c \begin{bmatrix} 1 \\ -k \\ k^2 \end{bmatrix}. \tag{42}$$

These columns are linearly dependent, implying rank deficiency and therefore structural non-identifiability.

5.9 Final Interpretation

Structural identifiability is not a property of data quality or estimation algorithms. It is a property of the model equations themselves.

If parameters only appear:

- as products,
- as ratios,
- or through unobserved state cascades,

then they do not constitute independent information directions in the output.

5.10 Key Takeaway

> *A parameter is structurally identifiable if and only if it generates a unique, independent direction in the output's Taylor expansion.*