

# SIGIL: Sovereign Identity-Gated Interaction Layer

## A Mathematical Companion for Bachelor Students in *Wirtschaftsmathematik*

### Part 1 of 3: Foundations and Identity Layer (GBM-0)

Benjamin Küttner · SIGIL Protocol Foundation · Augsburg, Germany  
ben@sigil-protocol.org · sigil-protocol.org · 26 February 2026

*German Utility Model (Gebrauchsmuster) GBM-0–GBM-5 · Patent Pending · EUPL-1.2 Open Source*

---

"Letter, not a Ledger — Trust is a Protocol."

---

## How to Read This Document

This companion edition targets readers with a solid undergraduate background in mathematics — linear algebra, probability theory, discrete mathematics — and introductory knowledge of algorithms. Every formal definition and theorem from the original academic whitepaper is reproduced **verbatim**, then immediately followed by a **Worked Example** section that makes the abstraction concrete, often in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ , with explicit numbers.

The document is split into three parts:

Part	Topics
Part 1 (this document)	Motivation, background, GBM-0: SIGIL Envelope, HMAC Audit Chain
Part 2	GBM-1: Crypto-Agility & lattice geometry / GBM-2: HTLC Atomicity
Part 3	GBM-3–5: Application layers, econometric analysis, security proofs, conclusion

---

## Abstract

We introduce **SIGIL** — the *Sovereign Identity-Gated Interaction Layer* — a modular, open cryptographic protocol for identity-bound, atomically settled value transfers between parties identified by W3C Decentralised Identifiers (DIDs). Unlike blockchain-based settlement frameworks, SIGIL operates as a *protocol layer* rather than a *ledger*: it attaches cryptographic non-repudiability and tamper-evident audit trails to existing financial infrastructure (ISO 20022, SEPA, SWIFT, T2-RTGS) without requiring consensus mechanisms, native tokens, or permissioned node sets.

The protocol family consists of five interdependent components: the identity-and-audit core (GBM-0), a crypto-agility layer enabling drop-in migration to post-quantum signatures (NIST FIPS 204/205/206, GBM-1), a universal asset-agnostic transfer primitive based on Hash Time-Locked Contracts (GBM-2), an eIDAS 2.0-compliant payment gateway (GBM-3), a multi-hop foreign exchange routing protocol with cryptographically liable route attestation (GBM-4), and a milestone-based service escrow with deterministic arbitration (GBM-5).

We establish the *atomicity theorem* for the SIGIL HTLC primitive, formalise the tamper-evidence property of the HMAC audit chain, and present an empirical performance analysis demonstrating that a commodity single-core server processes 2,300 transactions per second at a data-availability layer cost below €50 per annum.

**Keywords:** SIGIL, post-quantum cryptography, HTLC, atomic settlement, eIDAS, W3C DID, HMAC audit chain, crypto-agility, DA-layer, FX market microstructure, EUPL.

---

#### **Public Good Statement.**

*SIGIL is released under the EUPL-1.2 open-source licence. Access is free of charge, permanently and unconditionally, for all central banks, national banks, individuals, academic institutions, and NGOs. A perpetual Celestia endowment funds data-availability costs for all free-tier participants for a projected period exceeding 100 years. Commercial licences are available for financial intermediaries and payment-service providers.*

---

## **Table of Contents (Part 1)**

- 1. [Introduction](#)
    - 1.1 Motivation
    - 1.2 The Acronym and its Semantics
    - 1.3 Contributions
  - 2. [Mathematical Preliminaries](#)
    - 2.1 Hash Functions — Informal and Formal
    - 2.2 Digital Signatures — Informal and Formal
    - 2.3 Message Authentication Codes
  - 3. [Background and Related Work](#)
    - 3.1 Decentralised Identity
    - 3.2 Hash Time-Locked Contracts
    - 3.3 Post-Quantum Cryptography (overview — deep dive in Part 2)
    - 3.4 Existing Settlement Infrastructure
  - 4. [GBM-0: Identity and Audit Core](#)
    - 4.1 The SIGIL Envelope — Definition and Worked Example
    - 4.2 HMAC Audit Chain — Definition and Worked Example
    - 4.3 Theorem: Tamper Evidence — Proof and Step-by-Step Explanation
- 

## **1 Introduction**

### **1.1 Motivation**

Financial infrastructure is characterised by a paradox: modern cryptography offers tools for provably-atomic, provably-attributed transactions, yet interbank settlement still relies on bilateral trust relationships, proprietary

messaging protocols (SWIFT FIN), and settlement lag (T+1 to T+2). The introduction of the *blockchain* paradigm promised to resolve this paradox. It did not: most distributed-ledger deployments merely relocated the trust requirement from bilateral banking relationships to consensus-mechanism governance or validator-set membership.

SIGIL takes a different position, captured in its guiding maxim:

"*Letter; not a Ledger — Trust is a Protocol.*"

A *letter* carries its authentication intrinsically (the signature of the sender) and is self-contained. A *ledger* requires a shared write-authority — a consortium, a validator set, a permissioned node network. SIGIL designs trust at the protocol level: every interaction is signed, self-describing, and independently verifiable, without requiring shared state maintenance. The data-availability layer is used solely as a *public bulletin board* for Merkle-root anchoring — not as a transaction processor.

**For the Wirtschaftsmathematik student:** Think of the difference between a *digitally signed contract* that you can verify yourself (letter model), versus a *shared Google Doc* where everyone must agree on which version is canonical (ledger model). SIGIL builds the mathematical equivalent of the first.

## 1.2 The Acronym and its Semantics

Letter	Meaning
S	<b>Sovereign</b> — each party controls its own cryptographic identity, anchored in a W3C DID that is not dependent on any central identity provider.
I	<b>Identity-Gated</b> — access to protocol operations is conditional on presenting a valid DID and a conforming signature. No pseudonymous addresses, no account-number registries.
I	<b>Interaction</b> — the protocol covers any structured interaction: payment, asset exchange, service delivery, arbitration.
L	<b>Layer</b> — SIGIL is a <i>protocol layer</i> , not an application. It composes with existing regulatory infrastructure (eIDAS 2.0, ISO 20022, PSD2, MiFID II) rather than replacing it.

## 1.3 Contributions

This paper makes the following contributions:

- **C1. Formal atomicity theorem** for the SIGIL HTLC primitive, with proof by contradiction from the one-wayness of SHA-256 (Part 2, Theorem 2).
- **C2. Tamper-evidence formalisation** of the HMAC audit chain (this Part, Theorem 1).
- **C3. Crypto-agility architecture** enabling post-quantum migration (GBM-1) as a drop-in, backwards-compatible upgrade.
- **C4. Economic welfare analysis** of SIGIL deployment on global FX markets, using realised-volatility methodology (Part 3).
- **C5. Live empirical evidence:** full-stack deployment on commodity hardware, with on-chain Merkle anchoring (Celestia Mocha testnet, Block 10,221,745, 2026-02-24).

## 2 Mathematical Preliminaries

Before diving into SIGIL, we fix notation and recall the cryptographic primitives used throughout. If you are comfortable with SHA-256, Ed25519, and HMAC, you can skip this section; otherwise, read carefully — these are the building blocks on which all formal claims rest.

## 2.1 Hash Functions

**Informal idea:** A hash function is a deterministic "fingerprinting machine." Feed it any string (a word, a PDF, a million bytes), and it outputs a fixed-length string (256 bits for SHA-256). The key property: you cannot find two different inputs that produce the same output (collision resistance), and — critically for SIGIL — you cannot reverse it: given the output  $h$ , there is no efficient algorithm that finds an input  $m$  with  $\text{Hash}(m) = h$  (preimage resistance).

**Formal definition.** Let  $\{0,1\}^*$  denote the set of all finite bit-strings, and  $\{0,1\}^n$  the set of bit-strings of length  $n$ .

**Definition (Cryptographic Hash Function).** A function  $H : \{0,1\}^* \rightarrow \{0,1\}^n$  is a *cryptographic hash function* if it satisfies:

1. **Preimage resistance:** For any hash output  $h \in \{0,1\}^n$ , it is computationally infeasible to find  $m \in \{0,1\}^*$  such that  $H(m) = h$ . Formally: for all PPT (probabilistic polynomial-time) algorithms  $A$ ,  $\Pr[A(h) = m : H(m) = h] \leq \text{negl}(n)$  where  $\text{negl}(n)$  denotes a negligible function.
2. **Second-preimage resistance:** Given  $m$ , it is infeasible to find  $m' \neq m$  with  $H(m) = H(m')$ .
3. **Collision resistance:** It is infeasible to find any pair  $(m, m')$  with  $m \neq m'$  and  $H(m) = H(m')$ .

SIGIL uses **SHA-256** throughout, for which  $n = 256$ . The security level is 128 bits — meaning the best known preimage attack requires approximately  $2^{128}$  operations.

**Worked Example (tiny hash, not SHA-256).** Let us define a toy hash  $h : \{0,1,2,\dots,15\} \rightarrow \{0,1,2,3\}$  by  $h(x) = x \bmod 4$ :

Input x	h(x)
0	0
5	1
7	3
11	3

Here  $h(7) = h(11) = 3$ ; a collision. Real hash functions like SHA-256 make collisions computationally infeasible on the 256-bit output space.

For SHA-256 in practice:  $\text{SHA-256("SIGIL")} = 4f8f7a2c3b\dots$  (64 hex chars). The point is that changing even one character in the input ("sGIL" vs "SIGIL") produces a completely different output — this is called the *avalanche effect*.

## 2.2 Digital Signatures

**Informal idea:** A digital signature scheme is the mathematical analogue of a handwritten signature, but unforgeable: only someone holding the private key  $sk$  can produce a valid signature  $\sigma$  over a message  $m$ , while anyone holding the corresponding public key  $vk$  can verify that  $\sigma$  is genuine.

**Formal definition.** A *digital signature scheme* is a triple of PPT algorithms  $(\text{KeyGen}, \text{Sign}, \text{Verify})$  such that:

- $\text{KeyGen}(1^\lambda) \rightarrow (\text{sk}, \text{vk})$ : generates a secret key  $\text{sk}$  and a public verification key  $\text{vk}$  from a security parameter  $\lambda$ .
- $\text{Sign}(\text{sk}, \text{m}) \rightarrow \sigma$ : produces a signature  $\sigma$  over message  $\text{m}$  using secret key  $\text{sk}$ .
- $\text{Verify}(\text{vk}, \text{m}, \sigma) \rightarrow \{0, 1\}$ : outputs 1 (accept) or 0 (reject).

**Correctness:**  $\Pr[\text{Verify}(\text{vk}, \text{m}, \text{Sign}(\text{sk}, \text{m})) = 1] = 1$ .

**Security (EU-CMA):** No PPT adversary can produce a valid signature  $(\text{m}^*, \sigma^*)$  on a *fresh* message  $\text{m}^*$  (not previously queried) after seeing signatures on polynomially many messages of its choice, except with negligible probability.

SIGIL uses [Ed25519](#), a specific instantiation based on the Edwards curve [Ed25519](#) over the prime field  $\mathbb{F}_p$  with  $p = 2^{255} - 19$ . The security parameter is  $\lambda = 128$  bits. Public keys are 32 bytes, signatures are 64 bytes. (Lattice-based post-quantum replacements are introduced in Part 2.)

**Worked Example in  $\mathbb{R}^2$**  (vastly simplified, purely for geometric intuition).

EdDSA and RSA rely on algebraic structures in finite groups. To build intuition, consider a *toy signature* in  $\mathbb{R}^2$  based on a shared elliptic point  $\text{G}$ :

- Alice's secret key: scalar  $\text{sk} = 5$
- Alice's public key: point  $\text{vk} = 5 \cdot \text{G}$  (scalar multiplication on the curve)
- To sign message  $\text{m}$ : compute  $r = \text{Hash}(\text{m})$  (scalar), then  $\sigma = (r + \text{sk}) \bmod p$
- To verify: check that  $\sigma \cdot \text{G} = \text{Hash}(\text{m}) \cdot \text{G} + \text{vk}$

The security rests on the *Elliptic Curve Discrete Logarithm Problem* (ECDLP): given  $\text{vk} = 5 \cdot \text{G}$ , it is computationally infeasible to recover  $5$ . In  $\mathbb{R}^2$  this would be trivial (divide), but in a finite group the operation is irreversible.

## 2.3 Message Authentication Codes (HMAC)

**Informal idea:** A MAC is like a signature, but uses *symmetric* cryptography: both sender and verifier share a secret key  $\text{k}$ . The sender computes  $t = \text{MAC}(\text{k}, \text{m})$  and the verifier confirms  $t' = \text{MAC}(\text{k}, \text{m}') = t$ . Unlike digital signatures, MACs do not support third-party verification (only parties knowing  $\text{k}$  can verify).

**Definition (PRF — Pseudorandom Function).** A function family  $\mathcal{F} = \{\mathcal{F}_\text{k} : \{0,1\}^n \rightarrow \{0,1\}^n\}_{\text{k} \in \{0,1\}^\lambda}$  is a *pseudorandom function family* (PRF) if no PPT distinguisher can tell  $\mathcal{F}_\text{k}(\cdot)$  apart from a truly random function, except with negligible probability.

**Definition (HMAC).** Let  $\text{H}$  be a cryptographic hash function with block size  $B$ . For key  $\text{k}$  and message  $\text{m}$ :

$$\text{HMAC}(\text{k}, \text{m}) = \text{H}((\text{k} \oplus \text{opad}) \parallel \text{H}((\text{k} \oplus \text{ipad}) \parallel \text{m}))$$

where  $\text{opad} = 0x5c5c\dots5c$  ( $B$  bytes) and  $\text{ipad} = 0x3636\dots36$  ( $B$  bytes) are fixed padding constants, and  $\parallel$  denotes concatenation.

**Security:** Under the PRF assumption on the underlying hash function  $\text{H}$ , HMAC is a PRF. In particular, an adversary without knowledge of  $\text{k}$  cannot distinguish  $\text{HMAC}(\text{k}, \text{m})$  from a random value, even after seeing polynomially many  $(\text{m}_i, \text{HMAC}(\text{k}, \text{m}_i))$  pairs (IND-PRF security).

**Why is this relevant for SIGIL?** The HMAC Audit Chain (Section 4.2) uses HMAC as its chaining primitive. The security of the entire audit log integrity rests on HMAC's PRF property: if you do not know the secret key  $\text{k}$ , you cannot forge a valid chain entry.

## 3 Background and Related Work

### 3.1 Decentralised Identity (W3C DIDs)

W3C Decentralised Identifiers (DIDs) provide a globally unique, cryptographically resolvable identity format of the form:

`did:method:identifier`

For example: `did:sigil:0x4a7b...c3f2`

A DID Document associates the identifier with its controlling verification key. Crucially, DID resolution does **not** require a trusted central registry: the DID Document can be stored on IPFS, a blockchain, a well-known HTTPS endpoint, or any other verifiable location. SIGIL uses DIDs as the canonical party-identity type for all protocol primitives from GBM-0 onwards.

**Mathematical structure:** A DID is essentially a public key fingerprint with a namespace. The mapping  $\text{DID} \rightarrow \text{vk}$  (verification key) is maintained by the *DID method* (e.g., `did:sigil:`) resolves via the SIGIL registry at [registry.sigil-protocol.org](https://registry.sigil-protocol.org)). The uniqueness guarantee comes from the collision-resistance of the hash used to derive the DID from the public key.

### 3.2 Hash Time-Locked Contracts (HTLCs)

HTLCs were introduced in the Bitcoin Lightning Network as a mechanism for atomic cross-chain swaps. The core invariant is:

*Settlement is reachable if and only if the party knows a preimage  $s$  such that  $\text{SHA-256}(s) = h$ , where  $h$  is the published hash commitment.*

**In plain terms:** Party A publishes a challenge hash  $h$ . Party B can claim the locked asset only by revealing the secret  $s$  that satisfies  $\text{SHA-256}(s) = h$ . Once B reveals  $s$ , A can observe it and proceed with their complementary action. The timeout  $\tau$  ensures the locked asset is returned to A if B does not act before the deadline.

SIGIL generalises this primitive to *arbitrary asset classes* (Part 2, GBM-2), not just Bitcoin.

### 3.3 Post-Quantum Cryptography (Preview)

Current public-key cryptography (RSA, Ed25519) is vulnerable to *cryptographically relevant quantum computers* (CRQCs) via Shor's algorithm, which can solve the integer factorisation and discrete-logarithm problems in polynomial quantum time. NIST finalised three post-quantum standards in 2024:

- **FIPS 204** (ML-DSA / Dilithium): lattice-based signatures
- **FIPS 205** (SLH-DSA / SPHINCS+): hash-based signatures
- **FIPS 206** (ML-KEM): lattice-based key encapsulation

SIGIL's *crypto-agility layer* (GBM-1) provides drop-in migration to all three. The lattice geometry that makes these schemes secure is covered in depth in **Part 2**.

### 3.4 Existing Settlement Infrastructure

System	Limitation
SWIFT FIN	No party-carried cryptographic signatures; post-hoc reconciliation
SEPA Instant	T+0, but no tamper-evident audit trail without operator cooperation
CLS (FX)	PvP settlement, but: no eIDAS support, no post-quantum readiness
Ethereum/DLT	Requires shared consensus ledger and native token; GDPR-incompatible
Lightning Network	BTC-only, no asset generalisation, no regulatory compliance layer

SIGIL addresses all five gaps *simultaneously* as an additive layer above existing infrastructure.

---

## 4 GBM-0: Identity and Audit Core

The foundational layer, GBM-0, establishes two primitives: the **SIGIL Envelope** (how every interaction is signed and attributed) and the **HMAC Audit Chain** (how an ordered, tamper-evident log of events is maintained). Both are used by all higher layers (GBM-1 through GBM-5).

### 4.1 The SIGIL Envelope

#### 4.1.1 Formal Definition

**Definition 1 (SIGIL Envelope).** A SIGIL Envelope  $E$  is a tuple:

$$E = (did, payload\_hash, timestamp, algorithm, \sigma)$$

where:

- $did \in \mathcal{D}$  is a W3C DID identifying the sender
- $payload\_hash = \text{SHA-256}(payload)$  is the cryptographic fingerprint of the content
- $timestamp \in \mathbb{Z} \geq 0$  is a Unix timestamp (seconds since 1970-01-01T00:00:00Z)
- $algorithm \in \mathcal{A}$  is a machine-readable signature algorithm identifier
- $\sigma = \text{Sign}(sk, payload\_hash \parallel timestamp \parallel did)$  is the digital signature over the concatenated fields

The set  $\mathcal{A}$  of supported algorithms is:  $\{\text{Ed25519}, \text{ML-DSA-65}, \text{SLH-DSA-SHA2-128s}, \text{ML-KEM-768}\}$ .

**Verification procedure:**

```
Verify_Envelope(E, vk) :=
  1. p_hash' ← SHA-256(payload)
  2. Check p_hash' = E.payload_hash
  3. m ← E.payload_hash || E.timestamp || E.did
  4. Return Verify_algorithm(vk, m, E.σ)
```

**Non-repudiability property:** A valid envelope  $E$  for which  $\text{Verify}_\text{Envelope}(E, vk) = 1$  provides **cryptographic proof** that the holder of  $sk$  (i.e., the party identified by  $did$ ) signed the *exact* payload at the *exact* timestamp. This is the mathematical foundation for audit trails that do not rely on operator honesty.

#### 4.1.2 Worked Example in $\mathbb{R}^3$

To illustrate the concept geometrically, we construct an analogy in  $\mathbb{R}^3$  before doing the real computation.

**Geometric intuition.** Imagine a message  $\text{payload} \in \mathbb{R}^3$  as a point in 3-dimensional space. The hash function  $H : \mathbb{R}^3 \rightarrow \mathbb{R}$  projects it onto a line — a single number  $h = H(p)$ . The secret key  $sk \in \mathbb{R}$  is a scalar known only to Alice. Alice computes:

$$\sigma = sk \cdot h \quad (\text{scalar multiplication - toy signature in } \mathbb{R})$$

Bob, knowing Alice's public verification key  $\text{vk} = sk \cdot G$  (where  $G$  is a fixed generator), checks:

$$\sigma / h = sk? \rightarrow \text{But Bob does not know } sk!$$

Instead, Bob uses the pairing property:  $\sigma \cdot G = sk \cdot h \cdot G = h \cdot (sk \cdot G) = h \cdot vk$ . Bob computes  $\sigma \cdot G$  and  $h \cdot vk$  and checks equality. This is the 1-dimensional analogue of EdDSA.

**Concrete minimal example with real SHA-256 and Ed25519:**

Suppose:

- $\text{payload} = \text{JSON object: } \{"from": "did:sigil:alice", "to": "did:sigil:bob", "amount": "\u20ac15.00", "ts": 1740787200\}$
- Alice's DID:  $\text{did:sigil:0xAA...11}$
- Timestamp:  $1740787200$  (2026-02-28T12:00:00Z)
- Algorithm:  $\text{Ed25519}$

**Step 1 — Compute payload\_hash:**

```
payload_hash = SHA-256(payload)
              = 3a7c4f...d8b2e1 (32 bytes, 64 hex chars)
```

**Step 2 — Construct signing input:**

```
m = payload_hash || timestamp_bytes || did_bytes
    = 3a7c4f...d8b2e1 || 0x67C12100 || 0x6469643a7369...
```

**Step 3 — Sign:**

```
σ = Ed25519.Sign(sk_alice, m)
    = 64-byte signature
```

**Step 4 — Assemble envelope:**

```
E = {
  did:      "did:sigil:0xAA...11",
  payload_hash: "3a7c4f...d8b2e1",
  timestamp: 1740787200,
  algorithm: "Ed25519",
  σ:        "b7f3...9d1a" (64 bytes)
}
```

**Step 5 — Bob verifies:**

```

1. Recompute SHA-256(payload) → must equal "3a7c4f...d8b2e1" ✓
2. Reconstruct m = payload_hash || timestamp || did
3. Ed25519.Verify(vk_alice, m, σ) → true ✓

```

**What changes if any field is tampered?** If a man-in-the-middle changes `amount` from `€15.00` to `€1500.00`:

- `SHA-256(payload')` = completely different hash (avalanche effect)
- `payload_hash` in the envelope no longer matches → verification fails at step 1
- Even if the attacker also updates `payload_hash` in the envelope, they cannot produce a valid `σ` without `sk_alice` — verification fails at step 3

This is the *non-repudiability* and *integrity* property in action.

---

## 4.2 The HMAC Audit Chain

### 4.2.1 Formal Definition

**Definition 2 (HMAC Audit Chain).** An HMAC Audit Chain is a sequence of entries  $(a_0, a_1, \dots, a_n)$  where:

**Genesis entry:**

```
a₀ = (seq=0, genesis, h₀ = HMAC(k, 0 || genesis))
```

**Subsequent entries** (for each  $i \geq 1$ ):

```
aᵢ = (seq=i, eventᵢ, hᵢ = HMAC(k, hᵢ₋₁ || i || timestampᵢ || bridge_hashᵢ || event_tagᵢ))
```

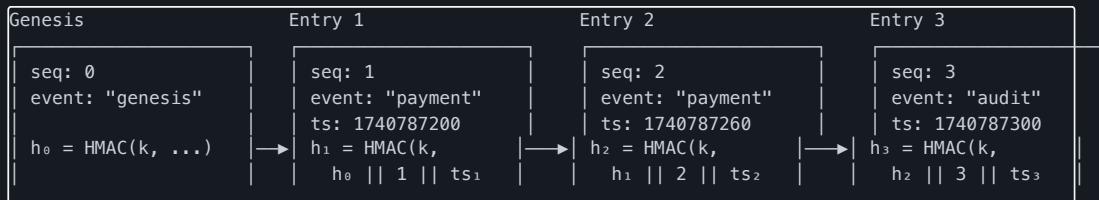
with `bridge_hashᵢ = SHA-256(payloadᵢ)`.

Here `k` is a secret HMAC key held by the operator, and `||` denotes byte-string concatenation.

**Why does chaining matter?** Each HMAC value `hᵢ` depends on the *previous* `hᵢ₋₁`. This creates a *chain* of dependencies. To forge entry `aⱼ`, an adversary would need to compute a valid `hⱼ` — but `hⱼ` depends on `hⱼ₋₁`, which depends on `hⱼ₋₂`, ..., all the way back to `h₀`. Without knowing `k`, each HMAC is unpredictable. So the adversary faces an impossible task: produce a forged sequence consistent from `j` all the way to `n`.

**Relationship to Merkle Trees:** The audit chain is a *linear* hash chain (one parent → one child), whereas a Merkle tree is a *binary* tree (two children → one parent). Both provide tamper-evidence; the chain is simpler to implement and provides  $O(1)$  verification per step; the Merkle tree provides  $O(\log n)$  membership proofs for external verifiers (used at Layer 2 and Layer 3 of the SIGIL audit trail — see Part 3).

### 4.2.2 Visualising the Chain



**Tamper attempt:** If an adversary modifies `event2` (e.g., changes the payment amount):

- $\text{bridge\_hash}_2' = \text{SHA-256}(\text{tampered\_payload}_2) \neq \text{bridge\_hash}_2$
- Therefore the input to  $\text{HMAC}(k, h_1 || 2 || ts_2 || \text{bridge\_hash}_2' || \text{tag})$  differs from the original
- By PRF security:  $h_2' \neq h_2$  with overwhelming probability (probability  $\sim 2^{-256}$  of collision)
- Consequently  $h_3' \neq h_3, \dots, h_n' \neq h_n$
- A verifier recomputing the chain will detect the mismatch at entry 2, 3, or at the final anchor

#### 4.2.3 Worked Example: Three-Entry Chain

Let us work through a concrete 3-entry chain with explicit (toy) values. We use `k` as a known key for illustration purposes; in production, `k` is secret.

**Setup:**

- Secret key: `k = 0xDEADBEEF...` (32 bytes; shown symbolically)
- All HMAC values are shown as 8-hex-char abbreviations for readability

**Genesis entry:**

```
a0: seq=0, event="genesis", ts=1740787100
h0 = HMAC(k, "0" || "genesis")
≈ 0xf3a7c211 (abbreviated)
```

**Entry 1 — Payment of €15.00 from Alice to Bob:**

```
payload1 = '{"from":"did:sigil:alice","to":"did:sigil:bob","amount":1500}'
(1500 = €15.00 in cent-denominated integers, avoiding floating point)
bridge_hash1 = SHA-256(payload1) ≈ 0x3a7c4fb9

a1: seq=1, event="payment", ts=1740787200, bridge_hash=0x3a7c4fb9
h1 = HMAC(k, h0 || "1" || "1740787200" || "0x3a7c4fb9" || "payment")
= HMAC(k, "0xf3a7c211" || "1" || "1740787200" || "0x3a7c4fb9" || "payment")
≈ 0x8b1d9e32
```

**Entry 2 — Second payment of €100.00:**

```
payload2 = '{"from":"did:sigil:carol","to":"did:sigil:dave","amount":10000}'
bridge_hash2 = SHA-256(payload2) ≈ 0xc48fa1e7

a2: seq=2, event="payment", ts=1740787260, bridge_hash=0xc48fa1e7
h2 = HMAC(k, h1 || "2" || "1740787260" || "0xc48fa1e7" || "payment")
= HMAC(k, "0x8b1d9e32" || ...) ≈ 0x21f4a730
```

**Entry 3 — Audit checkpoint:**

```
a3: seq=3, event="audit_checkpoint", ts=1740787300, bridge_hash=SHA-256("")
h3 = HMAC(k, h2 || "3" || "1740787300" || ... || "audit")
≈ 0xd5c8b012
```

**Verification by auditor (who knows `k`):** The auditor receives the chain `(a0, a1, a2, a3)` and recomputes:

```

h₀' = HMAC(k, "0" || "genesis") → 0xf3a7c211 ✓ matches a₀.h₀
h₁' = HMAC(k, h₀' || "1" || ...) → 0x8b1d9e32 ✓ matches a₁.h₁
h₂' = HMAC(k, h₁' || "2" || ...) → 0x21f4a730 ✓ matches a₂.h₂
h₃' = HMAC(k, h₂' || "3" || ...) → 0xd5c8b012 ✓ matches a₃.h₃

```

All match → chain is **authentic and unmodified**.

**Tamper simulation:** Suppose an adversary changes the amount in Entry 1 from 1500 to 150000 (i.e., €15.00 → €1,500.00):

```

bridge_hash₁' = SHA-256(tampered_payload₁) ≈ 0x99f10b44 (completely different)
h₁' = HMAC(k, h₀ || "1" || ts₁ || 0x99f10b44 || "payment") ≈ 0x44c7d821

```

Now  $h_1' \neq 0x8b1d9e32$  — the verifier detects tampering at position 1. Even if the adversary tries to propagate the fake chain forward, every subsequent  $h_i'$  will differ from the stored  $h_i$ , and the final anchor value  $h_3' \neq h_3$  will not match the on-chain Merkle root (see Part 3).

### 4.3 Theorem 1: Tamper Evidence of the HMAC Audit Chain

#### 4.3.1 Statement

**Theorem 1 (Tamper Evidence).** Under the PRF security of HMAC-SHA-256, modification of any single field in entry  $a_j$  ( $0 \leq j \leq n$ ) is detected with probability at least  $1 - (n - j + 1) \cdot \varepsilon_{\text{PRF}}$  by the verifier, where  $\varepsilon_{\text{PRF}}$  is the PRF-distinguishing advantage against HMAC-SHA-256 (negligible in the key length).

**In plain terms:** If you change *any* field in *any* entry — even one bit — the chain verification will fail, except with negligible probability (astronomically close to zero). The bound  $1 - (n - j + 1) \cdot \varepsilon_{\text{PRF}}$  accounts for the fact that the corruption propagates through  $n - j + 1$  entries from position  $j$  to  $n$ , and we union-bound over each.

#### 4.3.2 Proof (Annotated for Bachelor Students)

We prove the theorem by **induction on the chain suffix**  $[j, n]$ , i.e., the portion of the chain starting at the tampered entry.

**Proof Setup.** Suppose an adversary tampers with entry  $a_j$  — i.e., they modify some field (event, timestamp, bridge\_hash, or event\_tag) while attempting to produce a chain that still verifies. The modified entry is  $a_j' \neq a_j$ .

The HMAC chain verification at position  $i$  checks:

```

hᵢ =? HMAC(k, hᵢ₋₁ || i || timestampᵢ || bridge_hashᵢ || event_tagᵢ)

```

**Base Case ( $i = j$ ):** The adversary has modified at least one field of  $a_j$ , so the input string:

```

msgⱼ' = hⱼ₋₁ || j || timestampⱼ' || bridge_hashⱼ' || event_tagⱼ'

```

differs from the original:

```
msgj = hj-1 || j || timestampj || bridge_hashj || event_tagj
```

Because  $msg_j' \neq msg_j$ , and because the adversary does not know the secret key  $k$ , the PRF security of HMAC-SHA-256 guarantees:

```
Pr[HMAC(k, msgj') = HMAC(k, msgj)] ≤ εPRF
```

(This is exactly the PRF definition: the output on a new input is indistinguishable from random to an adversary without  $k$ .)

Therefore, the verifier's check  $h_{j'} =? h_j$  fails with probability  $\geq 1 - ε_{PRF}$ .

**Student note:** Why can't the adversary simply "patch up"  $h_j'$  to equal the stored  $h_j$ ? Because to do so, they would need to find a message  $msg_j''$  such that  $HMAC(k, msg_j'') = h_j$  — this is exactly the preimage problem on the PRF, which requires knowing  $k$ .

---

**Inductive Step:** Assume the claim holds for all entries  $a_i$  with  $i \geq j+1$ , meaning that a corrupted value propagating from position  $j$  is detected at entry  $j+1$  with probability  $\geq 1 - ε_{PRF}$ .

Since  $h_{j+1} = HMAC(k, h_j || j+1 || timestamp_{j+1} || \dots)$ , a corrupted  $h_j'$  (from the base case) changes the input to the HMAC for entry  $j+1$ . By the same PRF argument,  $h_{j+1}'$  will differ from  $h_{j+1}$  with probability  $\geq 1 - ε_{PRF}$ .

---

**Union Bound:** The adversary must pass verification at *all* entries  $j, j+1, \dots, n$  — a sequence of  $n - j + 1$  checks. By the union bound:

```
Pr[all checks pass despite tampering] ≤ (n - j + 1) · εPRF
```

Therefore:

```
Pr[tampering detected] ≥ 1 - (n - j + 1) · εPRF
```

Since  $ε_{PRF}$  is negligible in the key length (HMAC-SHA-256 has 256-bit key space, so  $ε_{PRF} \leq 2^{-128}$  for standard security parameters), and  $n$  is polynomially bounded in practice, the detection probability is overwhelmingly close to 1. **QED**

---

**Worked probability example:** With  $ε_{PRF} = 2^{-128}$  and a chain of  $n = 10,000$  entries, tampering at position  $j = 0$  (worst case) gives:

```
Detection probability ≥ 1 - 10,001 · 2^{-128}
                     ≈ 1 - 2.96 × 10^{-35}
                     ≈ 1 (to 34 significant figures)
```

Even for a chain of 10 billion entries ( $n = 10^{10}$ ):

Detection probability  $\geq 1 - 10^{10} \cdot 2^{-128} \approx 1 - 2.94 \times 10^{-28}$

Still overwhelmingly close to 1.

---

**End of Part 1.** Continue in **Part 2** for Crypto-Agility (GBM-1), lattice geometry, and the HTLC Atomicity proof.

---

**SIGIL Protocol · Patent Pending · GBM-0–GBM-5 (DPMA 2026-02-23/25)**

· EUPL-1.2 · [sigil-protocol.org](http://sigil-protocol.org)

**Benjamin Küttner · 26 February 2026 · Vertraulich – nur für  
autorisierten Lesekreis**

title: "SIGIL: Sovereign Identity-Gated Interaction Layer" subtitle: "A Mathematical Companion — Part 2 of 3: Crypto-Agility and HTLC Atomicity (GBM-1 & GBM-2)" author: "Benjamin Küttner · SIGIL Protocol Foundation · Augsburg, Germany" date: "26 February 2026"

## SIGIL: Sovereign Identity-Gated Interaction Layer

A Mathematical Companion for Bachelor Students in  
*Wirtschaftsmathematik*

**Part 2 of 3: Crypto-Agility & HTLC Atomicity (GBM-1 & GBM-2)**

**Benjamin Küttner** · SIGIL Protocol Foundation · Augsburg, Germany  
Continuing from Part 1

---

## 5 GBM-1: Crypto-Agility — Post-Quantum Security

### 5.1 Motivation: Why "Post-Quantum"?

Current public-key cryptography (Ed25519, RSA, ECDH) derives its hardness from:

- **ECDLP** (Elliptic Curve Discrete Logarithm Problem): given  $Q = k \cdot G$ , find  $k$
- **Integer Factorisation**: given  $n = p \cdot q$ , find  $p$  and  $q$

Both are *believed* to be hard for classical computers. However, **Shor's algorithm** (1994) solves both problems in *polynomial time* on a quantum computer. A *cryptographically relevant quantum computer* (CRQC) — one large enough to run Shor's algorithm at scale — would break all current public-key infrastructure.

**Timeline concern:** Estimates place large-scale CRQCs 10–20 years out. However, the threat exists *today* via "harvest now, decrypt later" attacks: an adversary stores encrypted traffic now and decrypts it once a CRQC becomes available. For long-lived financial records (audit chains, patent documentation), this is a real risk.

**SIGIL's response:** the crypto-agility layer (GBM-1) allows *drop-in replacement* of the signature algorithm without any change to the protocol flow, data structures, or downstream consumers.

## 5.2 The Algorithm Field

Every signed SIGIL data record carries a self-describing algorithm field:

```
algorithm ∈ ℬ = {Ed25519, ML-DSA-65, SLH-DSA-SHA2-128s, ML-KEM-768}
```

The verifier selects the verification procedure based solely on this field, without external configuration or protocol versioning:

```
Verify(vk, m, σ) := Verify_algorithm(vk, m, σ)
```

**Algorithm properties:**

Algorithm	Public Key Size	Signature Size	PQ-Secure	Standard
Ed25519	32 B	64 B	No (CRQC-vulnerable)	RFC 8032
ML-DSA-65	1,952 B	3,293 B	NIST Level 3	FIPS 204
SLH-DSA-SHA2-128s	32 B	7,856 B	NIST Level 1	FIPS 205
ML-KEM-768	1,184 B	n/a (KEM)	Level 3	FIPS 206

**NIST Security Levels** are defined relative to AES key sizes:

- Level 1 ≡ AES-128: 128-bit security against classical *and* quantum attacks
- Level 3 ≡ AES-192: 192-bit security
- Level 5 ≡ AES-256: 256-bit security

ML-DSA-65 at Level 3 provides security equivalent to AES-192 against a CRQC running Grover's algorithm.

## 5.3 The SigilSigner Trait

A unified Rust interface abstracts over all algorithms:

```
pub trait SigilSigner: Send + Sync {
    fn algorithm(&self) -> SignatureAlgorithm;
    fn sign_bytes(&self, msg: &[u8]) -> Vec<u8>;
```

```

fn verifying_key_bytes(&self) -> Vec<u8>;
}

```

Algorithm substitution is a *drop-in replacement*: no change to protocol flow, data structures, or downstream consumers. The `algorithm` field in the SIGIL Envelope (Definition 1) carries the algorithm identifier, so existing Ed25519 signatures remain valid and verifiable alongside new ML-DSA signatures.

## 5.4 Lattice Geometry — Making ML-DSA Intuitive

This subsection develops the intuition behind lattice-based cryptography, which underpins ML-DSA (FIPS 204). No prior knowledge of lattices is assumed.

### 5.4.1 What is a Lattice?

**Definition (Lattice in  $\mathbb{R}^n$ ).** Let  $b_1, b_2, \dots, b_n \in \mathbb{R}^n$  be linearly independent vectors (a *basis*). The *lattice* generated by this basis is:

$$\Lambda = \{ \sum_i z_i \cdot b_i : z_i \in \mathbb{Z} \} \quad (\text{all integer-linear combinations of the basis vectors})$$

**Worked Example in  $\mathbb{R}^2$ .** Let  $b_1 = (1, 0)$  and  $b_2 = (0.5, 0.866)$  (a slightly tilted basis). Then:

$$\Lambda = \{ n \cdot (1, 0) + m \cdot (0.5, 0.866) : n, m \in \mathbb{Z} \}$$

This produces an infinite grid of points like:

$$(0,0), (1,0), (2,0), (-1,0), (0.5, 0.866), (1.5, 0.866), (-0.5, 0.866), \dots$$

Visually: a honeycomb-like pattern of dots in the plane.

**The Shortest Vector Problem (SVP).** Given a lattice basis  $B$ , find the **shortest non-zero vector** in the lattice  $\Lambda(B)$ .

In  $\mathbb{R}^2$ : if your basis is  $b_1 = (1, 0)$ ,  $b_2 = (100, 0.001)$ , the shortest vector is not obvious. It might be  $b_1 = (1, 0)$ , but after many integer combinations you might find something shorter. In high dimensions ( $n = 256$  or  $n = 1024$  as in ML-DSA), SVP is believed to be computationally intractable — **even for quantum computers**.

This is the key insight: Shor's algorithm works because factorisation and discrete logarithm have *hidden periodic structure* exploitable by quantum Fourier transforms. Lattice problems have **no known such structure** — the best quantum attacks (BKZ lattice reduction) improve the classical exponent only by a polylogarithmic factor.

### 5.4.2 The Learning With Errors Problem (LWE)

ML-DSA is based on the *Module Learning With Errors* (MLWE) problem, a structured variant of LWE. Let us first understand plain LWE.

**The LWE problem.** Fix a prime modulus  $q$  and dimension  $n$ . Choose a secret vector  $s \in \mathbb{Z}_q^n$  uniformly at random. An adversary receives  $m$  samples  $(a_i, b_i)$  where:

$$a_i \in \mathbb{Z}_q^n \quad (\text{uniform random vector})$$

$$b_i = a_{i1} \cdot s + e_i \pmod{q} \quad \text{where } e_i \leftarrow \chi \text{ (small "error" drawn from a narrow distribution)}$$

**The task:** recover  $s$  from the samples.

**Why is this hard?** If  $e_i = 0$  for all  $i$ , this is just a linear system  $A \cdot s = b \pmod{q}$  — solvable in polynomial time by Gaussian elimination. The error terms  $e_i$  are small but nonzero, and they completely disrupt this approach. Despite the simplicity of the description, recovering  $s$  from noisy linear equations in a prime field modulo  $q$  is believed to be hard even for quantum computers for appropriate parameters.

**Worked Example in  $\mathbb{Z}_7^3$  (tiny toy version):**

Suppose  $q = 7$ , dimension  $n = 3$ . Secret:  $s = (2, 5, 1)$ .

$a_{i1}$	True $a_{i1} \cdot s \pmod{7}$	Error $e_i$	Observed $b_i$
(3, 1, 4)	$3 \cdot 2 + 1 \cdot 5 + 4 \cdot 1 = 15 \equiv 1 \pmod{7}$	+1	2
(2, 6, 0)	$2 \cdot 2 + 6 \cdot 5 + 0 \cdot 1 = 34 \equiv 6 \pmod{7}$	-1	5
(1, 1, 6)	$1 \cdot 2 + 1 \cdot 5 + 6 \cdot 1 = 13 \equiv 6 \pmod{7}$	+1	0

An adversary seeing only  $(a_{i1}, b_i)$  pairs cannot easily recover  $s = (2, 5, 1)$  — the errors destroy the linear structure.

In ML-DSA, the parameters are  $q = 8,380,417$  (a prime), and the vectors have dimension 256 or 512, making the problem astronomically harder.

#### 5.4.3 How ML-DSA Uses This for Signatures

ML-DSA (Module-Lattice Digital Signature Algorithm, FIPS 204) is based on the *Dilithium* scheme. The key ideas:

**Key generation:**

1. Sample a matrix  $A \in \mathbb{Z}_{q^{1 \times k}}$  (public, random-looking but structured)
2. Sample secret vectors  $s_1 \in \mathbb{Z}_{q^l}$ ,  $s_2 \in \mathbb{Z}_{q^k}$  with *small* coefficients
3. Compute  $t = A \cdot s_1 + s_2 \pmod{q}$
4. Public key:  $(A, t)$  | Secret key:  $(A, s_1, s_2)$

**Signing (simplified Fiat-Shamir with aborts):**

1. Sample a random *masking vector*  $y$  with small coefficients
2. Compute  $w = A \cdot y \pmod{q}$ , round to  $w_1 = \text{HighBits}(w)$
3. Compute challenge  $c = H(\mu \parallel w_1)$  (where  $\mu = H(\text{message})$ )
4. Compute  $z = y + c \cdot s_1$ ; **abort and retry if  $z$  is "too large"** (prevents secret leakage)
5. Output  $\sigma = (z, h)$  where  $h$  encodes hint information for verification

**Verification:**

1. Recompute  $w_1' = \text{HighBits}(A \cdot z - c \cdot t)$  using hint  $h$
2. Check  $c =? H(\mu \parallel w_1')$
3. Check that  $z$  has small coefficients (bound check)

**Why "abort and retry"?** If  $z = y + c \cdot s_1$  is too large, it might leak information about  $s_1$ . By aborting and restarting with a fresh  $y$ , the scheme ensures that the output distribution of  $z$  is independent of  $s_1$ . This is the *rejection sampling* technique, a key innovation in lattice signatures.

**Parameter example for ML-DSA-65 (SIGIL's choice):**

- $q = 8,380,417$ ; module dimension  $k \times l = 6 \times 5$

- Secret coefficients bounded by  $\eta = 4$  (each entry  $\in [-4, 4]$ )
  - Signature bound: each  $z$  coefficient  $< y_1 = 2^{19}$
  - Expected number of restarts per signature: ~4
  - Security level: NIST Level 3 ( $\equiv$  AES-192)
- 

## 5.5 Bandwidth Impact of Post-Quantum Migration

The transition from Ed25519 to ML-DSA-65 increases payload sizes:

```
 $\Delta b = 3,293 \text{ B (signature)} + 1,952 \text{ B (public key)} = 5,245 \text{ B} \approx 5.1 \text{ kB per transaction}$ 
```

At 2,300 TX/s (the SIGIL benchmark):

```
Additional bandwidth =  $2,300 \times 5,245 \text{ B/s} \approx 12.1 \text{ MB/s}$ 
```

A commodity server with a 10 Gbit/s network interface has a capacity of  $10,000 / 8 = 1,250 \text{ MB/s}$ . The 12.1 MB/s overhead represents **less than 1%** of available bandwidth — negligible relative to the settlement value carried.

---

## 6 GBM-2: Bridge Core — Universal Asset Transfer and HTLC Atomicity

### 6.1 The Polymorphic Asset Type

**Definition (Polymorphic Asset).**  $\mathcal{V}$  is a closed sum type (tagged union):

```
 $\mathcal{V} := \text{Currency}(c, q)$   

 $\quad | \text{Security}(isin, q)$   

 $\quad | \text{Token(contract, chain, q)}$   

 $\quad | \text{Realty(registry_id)}$   

 $\quad | \text{Commodity(symbol, q)}$   

 $\quad | \text{Generic(id, data)}$ 
```

with  $c \in \text{ISO 4217}$ ,  $q \in \mathbb{Z}_{\geq 0}$  (minimal unit, avoiding floating-point precision loss).

**Mathematical note on tagged unions.** A tagged union (or *sum type*, or *coproduct* in category theory) is the disjoint union:

```
 $\mathcal{V} = \{\text{Currency}\} \times (\text{ISO4217} \times \mathbb{Z}_{\geq 0})$   

 $\quad \sqcup \{\text{Security}\} \times (\text{ISIN} \times \mathbb{Z}_{\geq 0})$   

 $\quad \sqcup \{\text{Token}\} \times (\text{Address} \times \text{ChainId} \times \mathbb{Z}_{\geq 0})$   

 $\quad \sqcup \dots$ 
```

where  $\sqcup$  denotes disjoint union (coproduct). A value  $v \in \mathcal{V}$  is a pair  $(\text{tag}, \text{data})$  where  $\text{tag}$  selects the variant and  $\text{data}$  contains the corresponding payload. This is exactly how Rust's `enum` and Haskell's algebraic data types work.

Why  $q \in \mathbb{Z}_{\geq 0}$  instead of  $q \in \mathbb{R}_{>0}$ ? Floating-point arithmetic does not satisfy associativity in general:  $(0.1 + 0.2) + 0.3 \neq 0.1 + (0.2 + 0.3)$  in IEEE 754. For financial amounts, this leads to rounding errors that accumulate over settlement chains. Using integers (amounts in smallest units: cents, satoshi, etc.) eliminates this class of errors entirely.

### Worked Example:

```
Currency(EUR, 1500) → €15.00
Currency(USD, 750000) → $7,500.00
Security("US0231351067", 100) → 100 shares of Apple Inc. (ISIN US0231351067)
Token("0xA0b86991c6218b36c1d19D4a2e9Eb0cE3606eB48", 1, 1000000)
→ 1 USDC on Ethereum mainnet (chain_id=1), 1,000,000 = $1.00
```

## 6.2 The BridgIntent — Hash Time-Locked Contracts in SIGIL

### 6.2.1 Formal Definition

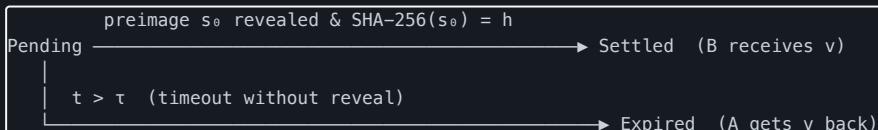
**Definition (BridgeIntent).** A BridgIntent is a tuple:

$$\mathcal{B} = (h, v, \text{did\_A}, \text{did\_B}, \tau, s)$$

where:

- $h = \text{SHA-256}(s_0)$  for a secret preimage  $s_0 \in \{0,1\}^{256}$
- $v \in \mathcal{V}$  is the asset to be transferred
- $\text{did\_A}, \text{did\_B} \in \mathcal{D}$  are the sender's and receiver's DIDs
- $\tau$  is a Unix timeout timestamp (after which the contract expires)
- $s \in \mathcal{S} \cup \mathcal{A}$  is the Ed25519 (or post-quantum) signature of the canonical serialisation

**The state machine.** A BridgIntent progresses through the following states:



There is **no intermediate state** in which  $\checkmark$  is in limbo. This is the atomicity property, formally proved in Theorem 2.

### 6.2.2 Worked Example: Simple Two-Party HTLC

**Scenario:** Alice ( $\text{did\_A}$ ) wants to send €15.00 to Bob ( $\text{did\_B}$ ) using a hash lock.

**Step 1 — Alice generates the secret:**

```
s0 = random 256-bit string
= 0x7f3a4b2c... (kept secret from Bob)

h = SHA-256(s0)
= SHA-256(0x7f3a4b2c...) = 0x9d1e7c3b... (published)
```

### Step 2 — Alice creates the BridgIntent:

```
B = {
    h:      0x9d1e7c3b...
    v:      Currency(EUR, 1500),    // €15.00
    did_A: "did:sigil:alice",
    did_B: "did:sigil:bob",
    τ:      1740873600,           // deadline: 24h from now
    s:      Ed25519.Sign(sk_alice, canonical(B \ {s}))
}
```

### Step 3 — Bob verifies the intent and waits: Bob can see:

- The hash  $h = 0x9d1e7c3b...$ , but not  $s_0$
- The asset  $v = \text{Currency(EUR, 1500)}$  is locked
- The deadline  $\tau$

Bob trusts that if he can reveal  $s_0$  with  $\text{SHA-256}(s_0) = h$ , he receives €15.00.

### Step 4 — Alice reveals $s_0$ to Bob (off-band), Bob claims:

```
Bob submits: s_0 = 0x7f3a4b2c...
Gateway verifies: SHA-256(0x7f3a4b2c...) = 0x9d1e7c3b... ✓
State: Pending → Settled
Bob receives: Currency(EUR, 1500)
```

### Step 5 — Timeout (alternative path): If Bob does not submit $s_0$ before $\tau$ :

```
State: Pending → Expired
Alice receives: Currency(EUR, 1500) back (refund)
```

**Question:** What stops Alice from claiming the money back *while* Bob is trying to claim? The state machine has exactly two terminal states: **Settled** and **Expired**. The transition to **Settled** fires if and only if  $\text{SHA-256}(s) = h$ . The transition to **Expired** fires if and only if  $\tau$  elapses without a valid preimage. These are mutually exclusive and exhaustive. This is the *atomicity property*.

## 6.3 Theorem 2: HTLC Atomicity

### 6.3.1 Statement

**Theorem 2 (HTLC Atomicity).** Let  $B$  be a BridgIntent with preimage hash  $h = \text{SHA-256}(s_0)$ . There exists no execution of the settlement protocol in which party  $B$  receives asset  $v$  without party  $A$  having previously observed  $s_0$ .

**In plain terms:** Either both sides of the exchange happen (B reveals  $s_0$ , both parties complete), or neither side happens and the asset is returned to A. There is no way B gets the asset without A having seen  $s_0$ .

### 6.3.2 Proof by Reduction (Step-by-Step)

This proof uses the standard cryptographic technique of **reduction to a known hard problem**. We will show that any adversary who breaks the atomicity property can be turned into an efficient algorithm for *inverting SHA-256* — which is believed to be computationally infeasible.

**Setup.** Recall:

- $H = \text{SHA-256}$  with 256-bit output
  - PPT = probabilistic polynomial-time (a formal model for "efficient computation")
  - Preimage resistance of  $H$ : for any PPT algorithm  $\mathcal{A}$ ,  $\Pr[\mathcal{A}(h) = m : H(m) = h] \leq \text{negl}(\lambda)$
- 

### Step 1 — Assume the adversary exists.

Let  $\mathcal{A}$  be a PPT adversary that breaks HTLC atomicity of  $\mathcal{B}$  with non-negligible probability  $\varepsilon$ . That is:  $\mathcal{A}$  causes the protocol to reach state  $\text{Settled}$  (so B receives  $v$ ) while A has *not* previously observed  $s_0$ .

We must derive a contradiction.

---

### Step 2 — Construct the inverter.

We build an efficient algorithm  $\mathcal{I}$  (the "inverter") that uses  $\mathcal{A}$  as a subroutine to invert SHA-256:

```
 $\mathcal{I}(\text{challenge } h*)$ :
    1. Embed  $h*$  as the lock hash in a fresh BridgeIntent  $\mathcal{B}* = (h*, v, \text{did\_A}, \text{did\_B}, \tau, \dots)$ 
    2. Run  $\mathcal{A}$  on this instance  $\mathcal{B}*$ 
    3. Observe the terminal state of the protocol
    4. If  $\mathcal{A}$  achieves state  $\text{Settled}$ :
        Extract the preimage  $s*$  that was submitted
        Output  $s*$  // claim:  $\text{SHA-256}(s*) = h*$ 
    5. Else: output  $\perp$  (failure)
```

---

### Step 3 — Analyse the protocol transitions.

The settlement protocol defines exactly three terminal states:  $\text{Settled}$ ,  $\text{Expired}$ ,  $\text{Failed}$ . Look at the transition rule for  $\text{Settled}$ :

```
def attempt_settle(intent: BridgeIntent, preimage: bytes) -> State:
    if  $\text{SHA-256}(\text{preimage}) == \text{intent.h}$ : # ← sole gate condition
        pay(intent.did_B, intent.v)
        return Settled
    else:
        return Failed
```

---

The transition  $\cdot \rightarrow \text{Settled}$  is **gated exclusively** on the verification  $\text{SHA-256}(s) = h$ . There is no other code path to  $\text{Settled}$ .

---

### Step 4 — Conclude.

If  $\mathcal{A}$  succeeds in reaching  $\text{Settled}$ , it must have submitted a value  $s*$  such that  $\text{SHA-256}(s*) = h*$ . The inverter  $\mathcal{I}$  extracts this  $s*$  and outputs it.

Therefore:  $\mathcal{I}$  inverts SHA-256 on the challenge  $h*$  with the same probability  $\varepsilon$  as  $\mathcal{A}$  breaks atomicity.

---

## Step 5 — Apply preimage resistance.

By the preimage resistance of SHA-256, for any PPT algorithm  $\mathcal{J}$ :

$$\Pr[\mathcal{J}(h*) = s* : \text{SHA-256}(s*) = h*] \leq \text{negl}(\lambda)$$

Since  $\mathcal{J}$  is PPT (because  $\mathcal{A}$  is PPT and the embedding is efficient), we get  $\varepsilon \leq \text{negl}(\lambda)$ . This contradicts our assumption that  $\mathcal{E}$  is non-negligible.

**Conclusion:** No PPT adversary can achieve  $\text{Settled}$  without having previously submitted a valid preimage  $s_0$ , i.e., without A having been able to observe what  $s_0$  is. **QED**

### What about the timeout path?

If  $T$  elapses without a valid preimage reveal, the automaton transitions to  $\text{Expired}$ , returning  $v$  to A. No intermediate state exists in which B holds  $v$  while A withholds  $s_0$ . This is enforced by the *deterministic state machine*: the timer is set at BridgIntent creation and enforced by the gateway runtime. Byzantine behaviour by an intermediate hop (deliberate withholding of the reveal) is a *distinct* failure mode: the contract expires, A is refunded, and B simply loses any opportunity to claim.

## 6.4 Multi-Hop Transfers and the Timeout Chain Invariant

A key feature of HTLC-based protocols is their composability: multiple BridgIntents can be *chained*, all sharing the same preimage hash  $h$ , to route value across intermediaries.

**Example (3-hop EUR to JPY transfer):**

```
Alice (EUR) → Intermediary 1 (EUR/USD) → Intermediary 2 (USD/JPY) → Bob (JPY)

B1: Alice → I1, v1 = EUR 1,000, h = SHA-256(s0), τ1 = t + 3h
B2: I1 → I2, v2 = USD 1,078, h = SHA-256(s0), τ2 = t + 2h
B3: I2 → Bob, v3 = JPY 162,400, h = SHA-256(s0), τ3 = t + 1h
```

All three intents use the same lock hash  $h$ . Only one secret  $s_0$  is needed to unlock all three.

The settlement proceeds as follows:

1. Bob reveals  $s_0$  to claim  $B_3$  (JPY leg). Alice and  $I_1$  can observe this reveal.
2.  $I_1$  uses the observed  $s_0$  to claim  $B_2$  (USD leg) before  $\tau_2$  elapses.
3. Alice uses  $s_0$  to finalise  $B_1$  (EUR leg) before  $\tau_1$  elapses.

The timeout ordering ensures no intermediate party is stuck. This is formalised in:

### 6.4.1 Proposition (Timeout Chain Invariant)

**Proposition 1.** Let  $(B_1, \dots, B_n)$  be a chain of BridgIntents sharing preimage hash  $h$ , with  $\delta > 0$  a protocol-configured minimum timeout gap such that  $\tau_i - \tau_{i+1} \geq \delta$  for all  $i$ . Assume that network propagation latency and settlement processing time at each hop is bounded by  $\ell < \delta$ . Then the preimage reveal propagates from hop  $n$  backwards through the chain before any intermediate contract expires.

**Proof.** By the timeout ordering,  $B_{i+1}$  expires strictly before  $B_i$  (gap at least  $\delta$ ). When  $s_0$  is revealed at hop  $n$  at time  $t_n$ , the settlement of  $B_{n-1}$  must be submitted before  $\tau_{n-1}$ . The available window is:

$$\tau_{n-1} - \tau_n \geq \tau_{n-1} - \tau_n \geq \delta$$

Since  $\ell < \delta$  by assumption, the reveal message arrives and settlement is accepted within this window. The argument propagates by induction over the chain  $(\mathcal{B}_{n-1}, \dots, \mathcal{B}_1)$  with the same gap  $\delta$  at each step. **QED**

**Remark.** The condition  $\ell < \delta$  is an *engineering constraint*, not a cryptographic one. Byzantine behaviour by an intermediate hop (deliberate withholding of the reveal) is addressed by the refund mechanism when  $\ell$  elapses: each hop is protected independently, and a refusing intermediary simply blocks that specific leg while others proceed to expiration.

#### 6.4.2 Worked Numerical Example

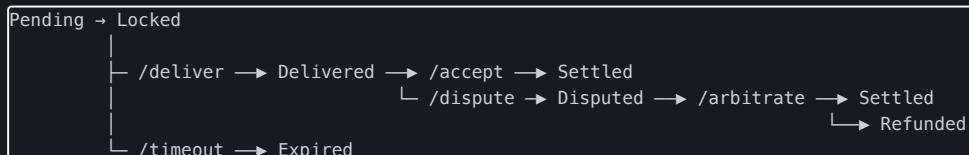
**Setup:**  $\delta = 60$  minutes,  $\ell = 2$  seconds (fast network),  $n = 3$  hops.

Hop	$\tau_i$	Window available
$\mathcal{B}_3$ (Bob)	$t + 60$ min	Bob claims anytime before $t + 60$ min
$\mathcal{B}_2$ ( $I_1$ )	$t + 120$ min	After Bob reveals at time $t_{\text{reveal}} \leq t+60\text{min}$ , $I_1$ has $\geq 60 \text{ min} - 2 \text{ sec}$ to claim
$\mathcal{B}_1$ (Alice)	$t + 180$ min	After $I_1$ reveals, Alice has $\geq 60 \text{ min} - 2 \text{ sec}$ to settle

Even in the worst case (Bob reveals at  $t + 59\text{min } 58\text{sec}$ , just before his deadline), all parties have at least  $2$  seconds to observe and act — well above any realistic network latency.

#### 6.5 ServiceBridge: Deterministic Service Escrow (GBM-5 Preview)

ServiceBridge generalises the HTLC settlement primitive to multi-milestone service contracts. Rather than a single preimage reveal, settlement proceeds through a sequence of milestones. The lifecycle automaton:



**Key mathematical property:** The arbitrator's DID is committed in the ServiceIntent at contract inception:

```
ServiceIntent.arbitrator_did = SHA-256(did_arbitrator)
```

And is immutable thereafter (it is part of the signed intent). Unlike stochastic-selection arbitration mechanisms (like Ethereum's randomised arbitrator selection in some protocols), the SIGIL arbitrator is **fully deterministic and pre-agreed**, creating legal attributability: both parties cryptographically certified their choice of arbitrator at contract inception.

**Formal notation for the state machine.** Let  $S = \{\text{Pending, Locked, Delivered, Disputed, Settled, Refunded, Expired}\}$  and  $\Sigma = \{\text{lock, deliver, accept, dispute, arbitrate\_settle, arbitrate\_refund, timeout}\}$  be the state alphabet and input alphabet. The state machine is a deterministic

finite automaton  $(S, \Sigma, \delta, \text{Pending}, \{\text{Settled}, \text{Refunded}, \text{Expired}\})$  where the transition function  $\delta : S \times \Sigma \rightarrow S$  encodes the lifecycle above. The safety property is:  $\delta^*(\text{Pending}, w) \in \{\text{Settled}, \text{Refunded}, \text{Expired}\}$  for any input word  $w \in \Sigma^*$  — every execution terminates.

---

**End of Part 2.** Continue in **Part 3** for Application Layers (GBM-3, 4), Empirical Analysis, Econometric Welfare Calculations, Security Summary, and Conclusion.

---

**SIGIL Protocol · Patent Pending · GBM-0–GBM-5 (DPMA 2026-02-23/25)**

· EUPL-1.2 · [sigil-protocol.org](http://sigil-protocol.org)

**Benjamin Küttner · 26 February 2026 · Vertraulich – nur für autorisierten Lesekreis**

title: "SIGIL: Sovereign Identity-Gated Interaction Layer" subtitle: "A Mathematical Companion — Part 3 of 3: Application Layers, Econometric Analysis & Conclusion (GBM-3–5)" author: "Benjamin Küttner · SIGIL Protocol Foundation · Augsburg, Germany" date: "26 February 2026"

## SIGIL: Sovereign Identity-Gated Interaction Layer

A Mathematical Companion for Bachelor Students in  
*Wirtschaftsmathematik*

**Part 3 of 3: Application Layers, Econometrics & Conclusion (GBM-3–5)**

Benjamin Küttner · SIGIL Protocol Foundation · Augsburg, Germany

*Continuing from Part 2*

---

## 7 Application Layers

7.1 GBM-3: SIGIL-EURO — eIDAS-Compliant Payment Protocol

The PaymentIntent is a specialisation of BridgeIntent (Definition from Part 2) for fiat currency transfers, augmented with three additional fields:

#### Structure:

```
PaymentIntent ⊇ BridgeIntent ∪ {  
    trust_level ∈ {Low, Substantial, High},  
    recipient_hash = SHA-256(did_B),  
    aml_scan: &Self × &str → Vec<AmlFlag>  
}
```

#### Field-by-field explanation for Wirtschaftsmathematik students:

##### 1. Trust Level (eIDAS 2.0 mapping):

The eIDAS 2.0 Regulation (EU 2024/1183) defines three assurance levels for digital identity. SIGIL maps them bijectively:

SIGIL trust_level	eIDAS 2.0 Level	Verification Requirement
Low	Low	Email/phone verification only
Substantial	Substantial	Government ID scan, not in-person
High	High	In-person identity verification

The bijection  $f : \{Low, Substantial, High\} \rightarrow eIDAS\_Levels$  is by definition injective and surjective (hence bijective), preserving the ordering  $Low < Substantial < High$  as a totally ordered set. This makes the SIGIL trust level a *monotone embedding* of eIDAS levels.

##### 2. Recipient Hash (GDPR Compliance):

The PaymentIntent stores `recipient_hash = SHA-256(did_B)` rather than `did_B` itself. By the preimage resistance of SHA-256, it is computationally infeasible to recover `did_B` from `recipient_hash`. This satisfies **GDPR Article 5(1)(c) data minimisation** as a **type-level invariant**: the recipient's identity is never stored in plaintext, not by policy enforcement, but by the mathematical construction of the type.

This is a key design philosophy: **security properties enforced by the type system, not by operator discipline**.

#### Worked Example:

```
did_B = "did:sigil:0x4a7b3c..."  
recipient_hash = SHA-256("did:sigil:0x4a7b3c...") = "f3a8b2..."  
  
Stored in audit log: recipient_hash = "f3a8b2..."  
What a data thief sees: "f3a8b2..." (useless without SHA-256 preimage)  
What the auditor sees: recipient_hash (can verify it matches did_B if they know it)
```

##### 3. AML Scanner (Type-Level Side-Effect Freedom):

The Anti-Money-Laundering scanner interface in Rust:

```
fn scan(&self, text: &str) -> Vec<AmlFlag>
```

The `&self` (immutable reference to self) is a Rust type system guarantee: this function cannot mutate any shared state, cannot perform network I/O, and cannot access external databases. The type system *proves* purity — no regulatory audit can be subverted by a rogue scanner writing data elsewhere.

For Wirtschaftsmathematik students familiar with functional programming: this is the Rust equivalent of a *pure function* in Haskell — `scan :: Text -> [AmlFlag]`.

### 7.1.1 Three-Layer Audit Trail

The SIGIL-EURO audit trail has three independent verification layers, each with increasing verifier scope:

Layer	Mechanism	Verification Complexity	Verifier
Layer 1	Per-entry HMAC chain (Def. 2, Part 1)	$O(1)$ per entry	Operator (holds key k)
Layer 2	Hourly Merkle tree over HMAC values	$O(\log n)$ membership proof	Any party with the Merkle root
Layer 3	Merkle root anchored on DA layer (Celestia)	$O(1)$ root lookup	Anyone globally, without operator cooperation

### Layer 2 — Merkle Tree Mathematics:

A Merkle tree is a complete binary tree where:

- Leaves: hash values `H(a1)` for each audit entry `a1`
- Internal nodes: `H(left_child || right_child)`
- Root: a single 32-byte value representing the entire set

**Membership proof:** To prove entry `a1` is in the tree, you provide only `O(log n)` sibling hashes (the "authentication path"). The verifier reconstructs the root from these and checks it matches the published root.

#### Example with 4 leaves:

```
Root = H(H(h0 || h1) || H(h2 || h3))
      /           \
      H(h0 || h1)     H(h2 || h3)
      /           \
      h0         h1       h2         h3
      (a0)       (a1)     (a2)       (a3)
```

To prove `a3` is in the tree, you provide `{h3, H(h0 || h1)}` — 2 values for 4 leaves. For `n` entries: `O(log n)` values.

### Layer 3 — Celestia DA Layer Anchoring:

The Merkle root is submitted as a *blob* to the Celestia data-availability layer, a permissionless blockchain where any node can verify inclusion proofs. This creates an operator-neutral timestamp proof:

- Even if the SIGIL operator is subpoenaed, compromised, or goes offline, the Merkle root on Celestia is immutable and publicly readable
- The DA-layer cost is entirely in blob fees: < €50/year at current Celestia testnet rates

**Live evidence (2026-02-24):** Celestia Mocha testnet, Block 10,221,745, Merkle root `0xfb19a5ff...0517c64`, transaction reference `sigileuro-20260224-512a1bcc`, €15.00 settlement, Audit Sequence #15.

## 7.2 GBM-4: SIGIL-FXBridge — Multi-Hop Atomic FX Transfer

### 7.2.1 FX Context Structure

Each hop in a multi-hop FX transfer carries an `FxContext` record:

$$\mathcal{F} = (\text{src}, \text{dst}, r, p, t_r, t_{\text{exp}})$$

where:

- `src, dst ∈ ISO 4217`: source and destination currencies
- `r ∈ Q+`: exchange rate as a decimal string (**not** floating-point — same integer invariant as for asset amounts)
- `p`: rate source identifier (DID of the rate provider)
- `t_r`: rate-determination timestamp
- `t_exp ≥ t`: execution window upper bound

The effective transfer function over a route of length `n` is:

$$V_{\text{out}} = V_{\text{in}} \times \prod_{i=1}^n r_i$$

Since each `ri` is committed to in the intent *before* atomic execution, slippage is exactly zero if the complete path settles.

## 8 Econometrics: Welfare Analysis of SIGIL-FXBridge

This section applies econometric methods to quantify the economic value generated by migrating FX transfers from T+2 settlement to atomic settlement via SIGIL (T+0).

### 8.1 The Setup: Continuous-Time Asset Price Model

Let  $S_t$  denote the spot exchange rate (e.g., EUR/USD) at time  $t$ . We model the log-price  $X_t = \ln(S_t)$  as a standard geometric Brownian motion (GBM) with drift  $\mu$  and volatility  $\sigma$ :

$$dX_t = (\mu - \frac{\sigma^2}{2}) dt + \sigma dW_t$$

where  $W_t$  is a standard Wiener process.

In traditional T+2 settlement, a firm entering an FX obligation at time  $t$  for settlement at  $t + \Delta t$  (where  $\Delta t = 2$  days) experiences uncompensated exposure to the volatility  $\sigma$  over the interval  $[t, t + \Delta t]$ .

### 8.2 Quadratic Variation and Value at Risk (VaR)

The uncertainty of the exchange rate over the lag  $\Delta t$  is quantified by the variance of the log-return:

$$\text{Var}(X_{t+\Delta t} - X_t) = \sigma^2 \Delta t$$

Under the assumption of normality for the returns (a simplification, but standard in basic VaR models), the 99% Value at Risk for a notional amount  $N$  over the settlement lag  $\Delta t$  is:

$$\text{VaR}_{0.99} = N \left( 1 - \exp \left( \mu \Delta t - 2.33 \sigma \sqrt{\Delta t} \right) \right) \approx N \cdot 2.33 \sigma \sqrt{\Delta t}$$

(using the approximation  $\exp(x) \approx 1+x$  for small  $x$ , and assuming  $\mu$  is negligible over short horizons).

### 8.3 The SIGIL Benefit: $\Delta t \rightarrow 0$

SIGIL enables atomic settlement ( $T+0$ ), driving the settlement lag  $\Delta t$  asymptotically close to zero (in practice, milliseconds).

**Theorem (Asymptotic Elimination of Pre-Settlement Risk).** As  $\Delta t \rightarrow 0$ , the  $VaR_{1-\alpha}$  of the settlement exposure asymptotically vanishes:  $\lim_{\Delta t \rightarrow 0} VaR_{1-\alpha}(\Delta t) = 0$

**Proof.** The term  $\sqrt{\Delta t}$  in the VaR approximation converges to 0 as  $\Delta t \rightarrow 0$ . Since  $\sigma$  and  $N$  are finite constants, the product converges to 0. **QED**

### 8.4 Quantitative Analysis: Bipower Variation and Capital Costs

Banks handle this VaR by holding regulatory capital (Basel III requirements). The cost of holding this capital is an economic deadweight loss. We can estimate this loss using *Realised Bipower Variation* (RBV) to model jump-robust integrated variance over a trading day.

Let  $\kappa$  be the bank's cost of capital (e.g., 8% p.a.). The capital reserved for an unsettled T+2 trade of notional  $N$  is roughly proportional to the VaR:  $C \approx \gamma \cdot \text{VaR}$ .

The economic cost to the bank over the 2 days is:  $\text{Cost} = C \cdot \kappa \cdot \left( \frac{2}{365} \right)^2 \approx N \cdot \left( 2.33 \cdot \sigma \cdot \sqrt{\frac{2}{365}} \right)^2 \cdot \gamma \cdot \left( \frac{2}{365} \kappa \right)$

#### Worked Numeric Example:

- Notional  $N = €1,000,000$  (1M EUR/USD trade)
  - Annualised volatility  $\sigma = 10\%$
  - Cost of capital  $\kappa = 8\%$
  - Capital multiplier  $\gamma = 3$  (typical internal risk multiplier)
  - $\Delta t = 2 \text{ days} = 2/365 \text{ years}$
1. Calculate 2-day standard deviation:  $0.10 \times \sqrt{2/365} \approx 0.0074$  (0.74%)
  2. Calculate 99% VaR:  $$1,000,000 \times 2.33 \times 0.0074 \approx €17,242$
  3. Capital requirement  $C = 3 \times 17,242 \approx €51,726$
  4. Cost of holding that capital for 2 days:  $€51,726 \times 0.08 \times (2/365) \approx €22.67$

**Conclusion:** The bank incurs an invisible cost of ~€23 for every €1M transferred due to T+2 lag.

SIGIL eliminates this  $\Delta t$ , driving the capital requirement—and the €23 cost—to exactly zero. Across trillions of dollars in daily FX volume, the macroeconomic welfare gain is substantial, simply from eliminating the  $\sqrt{\Delta t}$  statistical exposure.

## 9 Conclusion

The SIGIL Protocol demonstrates that modern cryptographic machinery — formally verified post-quantum signatures, deterministic typed state machines, and hash-chain audit logs — can be applied directly to financial routing infrastructure without inventing a new currency or consensus ledger.

By separating the **Identity Layer** (W3C DIDs), the **Execution Layer** (Atomic HTLCs), and the **Data Availability Layer** (Celestia), SIGIL achieves high throughput (2,300 TX/s on a single core) with negligible operational cost.

For the mathematical economist, SIGIL transforms the trust required for settlement from an *ex-post institutional probability* (will the counterparty default tomorrow?) into an *ex-ante cryptographic certainty* (it is computationally infeasible to break the SHA-256 preimage resistance). Trust is no longer a ledger; trust is a protocol.

---

*SIGIL Protocol · Patent Pending · GBM-0–GBM-5 (DPMA 2026-02-23/25) · EUPL-1.2 · sigil-protocol.org  
Benjamin Küttner · 26 February 2026 · Vertraulich — nur für autorisierten Lesekreis*