Euler-Maclaurin formula and generalised iterated integrals

Carlo Bellingeri

With
P. Friz (TU Berlin & WIAS)
S. Paycha (Potsdam University)

Algebraic, analytic and geometric structures emerging from quantum field theory
11.03.2024

The Euler-Maclaurin Formula

For any $f : [0, N] \to \mathbb{R}$ of class C^p we have

$$(EML) \quad \sum_{k=0}^{N-1} f(k) = \int_0^N f(s)\mathrm{d}s + \sum_{k=1}^{p} \frac{B_k}{k!}[f^{(k-1)}(s)]_{s=0}^N + R_p$$

B_k Bernoulli numbers $1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \cdots$

R_p Remainder which depends linearly on $f^{(p)}$.

This identity links integrals to sums, leading to many results in mathematics: complex analysis, operator theory, numerical analysis.

What happens in a stochastic framework with Riemann-Stieltjes integration?

The Euler-Maclaurin Formula

For any $f : [0, N] \to \mathbb{R}$ of class C^p we have

$$(EML) \quad \sum_{k=0}^{N-1} f(k) = \int_0^N f(s)\mathrm{d}s + \sum_{k=1}^{p} \frac{B_k}{k!}[f^{(k-1)}(s)]_{s=0}^{N} + R_p$$

B_k Bernoulli numbers $1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \cdots$

R_p Remainder which depends linearly on $f^{(p)}$.

This identity links integrals to sums, leading to many results in mathematics: complex analysis, operator theory, numerical analysis.

What happens in a stochastic framework with Riemann-Stieltjes integration?

The Euler-Maclaurin Formula

For any $f : [0, N] \to \mathbb{R}$ of class C^p we have

$$(EML) \quad \sum_{k=0}^{N-1} f(k) = \int_0^N f(s)\mathrm{d}s + \sum_{k=1}^{p} \frac{B_k}{k!}[f^{(k-1)}(s)]_{s=0}^N + R_p$$

B_k Bernoulli numbers $1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \cdots$

R_p Remainder which depends linearly on $f^{(p)}$.

This identity links integrals to sums, leading to many results in mathematics: complex analysis, operator theory, numerical analysis.

What happens in a stochastic framework with Riemann-Stieltjes integration?

(EML) is justified on the basis of three basic identities

- The sum $\sum$ and the integration $\int$ are the inverse operators of

$$\delta f(x) = f(x+1) - f(x), \quad Df(x) = f'(x).$$

- The formal expansion

$$\delta f(x) = \sum_{k\geq 1} \frac{f^{(k)}(x)}{k!} = \sum_{k\geq 1} \frac{D^k}{k!} f(x) = (e^D - \text{id}) f(x)$$

- The series expansion

$$\frac{z}{e^z - 1} = \sum_{k\geq 0} \frac{B_k}{k!} z^k$$

(EML) is justified on the basis of three basic identities

- The sum $\sum$ and the integration $\int$ are the inverse operators of

$$\delta f(x) = f(x+1) - f(x)\,, \quad Df(x) = f'(x)\,.$$

- The formal expansion

$$\delta f(x) = \sum_{k\geq 1} \frac{f^{(k)}(x)}{k!} = \sum_{k\geq 1} \frac{D^k}{k!} f(x) = (e^D - \mathrm{id}) f(x)$$

- The series expansion

$$\frac{z}{e^z - 1} = \sum_{k\geq 0} \frac{B_k}{k!} z^k$$

The Euler-Maclaurin Formula

(EML) is justified on the basis of three basic identities

- The sum $\sum$ and the integration $\int$ are the inverse operators of

$$\delta f(x) = f(x+1) - f(x)\,, \quad Df(x) = f'(x)\,.$$

- The formal expansion

$$\delta f(x) = \sum_{k\geq 1} \frac{f^{(k)}(x)}{k!} = \sum_{k\geq 1} \frac{D^k}{k!} f(x) = (e^D - \mathrm{id}) f(x)$$

- The series expansion

$$\frac{z}{e^z - 1} = \sum_{k\geq 0} \frac{B_k}{k!} z^k$$

The Euler-Maclaurin Formula

(EML) is justified on the basis of three basic identities

- The sum $\sum$ and the integration $\int$ are the inverse operators of

$$\delta f(x) = f(x+1) - f(x)\,, \quad Df(x) = f'(x)\,.$$

- The formal expansion

$$\delta f(x) = \sum_{k\geq 1} \frac{f^{(k)}(x)}{k!} = \sum_{k\geq 1} \frac{D^k}{k!} f(x) = (e^D - \mathrm{id}) f(x)$$

- The series expansion

$$\frac{z}{e^z - 1} = \sum_{k\geq 0} \frac{B_k}{k!} z^k$$

By combining these three facts

$$\sum = \delta^{-1} = (e^D - \mathrm{id})^{-1} = D^{-1} \frac{D}{e^D - \mathrm{id}}$$
$$= D^{-1} + \sum_{k \geq 1} \frac{B_k}{k!} D^{k-1} = \int + \sum_{k \geq 1} \frac{B_k}{k!} D^{k-1}$$

This approach does not take account of boundary terms.

Using this approach, classical sums can be replaced by a generic sum over the vertices of a polytope [Berline-Vergne '18] and other generalizations

By combining these three facts

$$\sum = \delta^{-1} = (e^D - \mathrm{id})^{-1} = D^{-1}\frac{D}{e^D - \mathrm{id}}$$
$$= D^{-1} + \sum_{k\geq 1}\frac{B_k}{k!}D^{k-1} = \int + \sum_{k\geq 1}\frac{B_k}{k!}D^{k-1}$$

This approach does not take account of boundary terms.

Using this approach, classical sums can be replaced by a generic sum over the vertices of a polytope [Berline-Vergne '18] and other generalizations

An alternative proof

The same formula can be proved as a specific case of a family of EML-type formulas [Boas '77]

The key point is to rewrite the sum as a Stieltjes integral.

$$\sum_{k=0}^{N-1} f(k) = \int_{-a}^{N-a} f(s)\mathrm{d}[s] \quad 0 < a < 1$$

Applying an integration by parts with the function $P_1(s) = s - [s]$

$$\sum - \int = - \int_{-a}^{N-a} f(s)\mathrm{d}(P_1(s))$$

$$= [-f(s)P_1(s)]_{-a}^{N-a} + \int_{-a}^{N-a} f'(s)P_1(s)\mathrm{d}s$$

What happens if we iterate the integration?

An alternative proof

The same formula can be proved as a specific case of a family of EML-type formulas [Boas '77]

The key point is to rewrite the sum as a Stieltjes integral.

$$\sum_{k=0}^{N-1} f(k) = \int_{-a}^{N-a} f(s)\mathrm{d}[s] \quad 0 < a < 1$$

Applying an integration by parts with the function $P_1(s) = s - [s]$

$$\sum - \int = - \int_{-a}^{N-a} f(s)\mathrm{d}(P_1(s))$$

$$= [-f(s)P_1(s)]_{-a}^{N-a} + \int_{-a}^{N-a} f'(s)P_1(s)\mathrm{d}s$$

What happens if we iterate the integration?

An alternative proof

For any sequence of functions $(P_k)_{k\geq 1}$ s.t. $P_k' = P_{k-1}$ we have

$$\sum = \int + \sum_{l=1}^{p} (-1)^l [f^{(l-1)}(s)P_l(s)]_{-a}^{N-a} + (-1)^{p+1} \int_{-a}^{N-a} f^{(p)}(s)P_p(s)\mathrm{d}s$$

Sending $a \downarrow 0$, one has EML formula modulo initial constants

The classic choice is to consider $P = (Q_k)_{k\geq 1}$ s.t.

$$\begin{cases} Q_k' = Q_{k-1} \\ Q_1 = t - [t] - \frac{1}{2} \\ Q_k \quad \text{1-periodic}, \int_0^1 Q_k(s)\mathrm{d}s = 0 \end{cases}$$

The system identifies a unique solution such that $Q_k(0) = \frac{B_k}{k!}$.
Functions Q_k are called periodic Bernoulli polynomials.

An alternative proof

For any sequence of functions $(P_k)_{k\geq 1}$ s.t. $P_k' = P_{k-1}$ we have

$$\sum = \int + \sum_{l=1}^{p} (-1)^l [f^{(l-1)}(s)P_l(s)]_{-a}^{N-a} + (-1)^{p+1} \int_{-a}^{N-a} f^{(p)}(s) P_p(s) \mathrm{d}s$$

Sending $a \downarrow 0$, one has EML formula modulo initial constants

The classic choice is to consider $P = (Q_k)_{k\geq 1}$ s.t.

$$\begin{cases} Q_k' = Q_{k-1} \\ Q_1 = t - [t] - \frac{1}{2} \\ Q_k \quad \text{1-periodic}\,, \int_0^1 Q_k(s)\mathrm{d}s = 0 \end{cases}$$

The system identifies a unique solution such that $Q_k(0) = \frac{B_k}{k!}$.
Functions Q_k are called periodic Bernoulli polynomials.

Extension to Stieltjes Integrals

Let $X\colon [0,N] \to V$ be a finite variation process and $f\colon V \to L(V,W)$ a smooth function with V, W Banach spaces. We consider the Riemann-Stieltjes integral.

$$\int_0^N f(X_s)dX_s$$

we want to compare it with Riemann-Stieltjes sums

$$I_N(f,X) := \sum_{k=0}^{N-1} f(X_k)\delta X_k\,, \quad \delta X_k = X_{k+1} - X_k$$

Is there a way of deriving a formula between these two operators?
Yes, but we must use the Boas approach.

Let $X\colon [0,N] \to V$ be a finite variation process and $f\colon V \to L(V,W)$ a smooth function with V, W Banach spaces. We consider the Riemann-Stieltjes integral.

$$\int_0^N f(X_s)dX_s$$

we want to compare it with Riemann-Stieltjes sums

$$I_N(f,X) := \sum_{k=0}^{N-1} f(X_k)\delta X_k\,, \quad \delta X_k = X_{k+1} - X_k$$

Is there a way of deriving a formula between these two operators? Yes, but we must use the Boas approach.

Extension to Stieltjes Integrals

Thanks to the finite variation we write

$$\sum_{k=0}^{N-1} f(X_k)(X_{k+1} - X_k) = \int_{-a}^{N-a} f(X_s)\mathrm{d}X_{[s+1]}$$

Using the notation $Z_s^1 = X_s - X_{[s+1]}$ we repeat

$$I_N(f, X) - \int_{-a}^{N-a} f(X_s)dX_s = -\int_{-a}^{N-a} f(X_s)\mathrm{d}(Z_s^1 + b^1)$$

$$= -[f(X_s)(Z_s^1 + b^1)]_{-a}^{N-a} + \int_{-a}^{N-a} Df(X_s)(\mathrm{d}X_s \otimes Z_s^1 + b^1)ds$$

By iterating this integration by parts, we obtain a sequence of non-linear path functionals on X.

Thanks to the finite variation we write

$$\sum_{k=0}^{N-1} f(X_k)(X_{k+1} - X_k) = \int_{-a}^{N-a} f(X_s)\mathrm{d}X_{[s+1]}$$

Using the notation $Z_s^1 = X_s - X_{[s+1]}$ we repeat

$$I_N(f, X) - \int_{-a}^{N-a} f(X_s)dX_s = -\int_{-a}^{N-a} f(X_s)\mathrm{d}(Z_s^1 + b^1)$$
$$= -[f(X_s)(Z_s^1 + b^1)]_{-a}^{N-a} + \int_{-a}^{N-a} Df(X_s)(\mathrm{d}X_s \otimes Z_s^1 + b^1)ds$$

By iterating this integration by parts, we obtain a sequence of non-linear path functionals on X.

Tensor notation

For each $f\colon V \to L(V,W)$ we write the gradient

$$D^p f\colon V \to L(V^{\otimes p+1}, W)$$

To keep track of higher order tensors, we use unitary tensor space

$$T_1((V)) = \mathbf{1} \oplus \prod_{i=1} V^{\otimes i}, \quad T_1^p(V) = \mathbf{1} \oplus \bigoplus_{i=1}^{p} V^{\otimes i}$$

with its applications of projections $\pi_i\colon T_1((V)) \to V^{\otimes i}$ and its inverse application $v \to v^{-1}$.

For each $f\colon V \to L(V, W)$ we write the gradient

$$D^p f\colon V \to L(V^{\otimes p+1}, W)$$

To keep track of higher order tensors, we use unitary tensor space

$$T_1((V)) = \mathbf{1} \oplus \prod_{i=1} V^{\otimes i}, \quad T_1^p(V) = \mathbf{1} \oplus \bigoplus_{i=1}^{p} V^{\otimes i}$$

with its applications of projections $\pi_i\colon T_1((V)) \to V^{\otimes i}$ and its inverse application $v \to v^{-1}$.

Sawtooth Signature

Definition (B.,Friz, Paycha '24+)

For any $b \in T_1((V))$, *we consider the finite variation process* $Z(b)\colon [0, N] \to T_1((V))$ *given by*

$$\pi_1(Z_t(b)) = X_t - X_{[t+1]} + \pi_1(b)$$

and the higher order components satisfy for all $k \geq 2$ *the differential equations*

$$\begin{cases} d\pi_k(Z_t(b)) = dX_t \otimes \pi_{k-1}(Z_t(b)) \\ Z_0(b) = \pi_k(b) \end{cases}$$

We call $Z(b)$ *the sawtooth signature with initial data* b.

Preliminary EML formula

Proposition (B.,Friz, Paycha '24+)

For any $b \in T_1^p(V)$

$$I_N(f,X) = \int_0^N f(Y_t)\mathrm{d}X_t - [f(X_t)b^1]_{t=0}^N + \sum_{l=2}^{p}(-1)^l [D^{l-1}f(X_t)\pi_l(Z_t(b))]_{t=0}^N + R^p(b)$$

where the remainder $R^m(b)$ *is given by the Stieltjes integral*

$$R^p(b) = (-1)^{p+1}\int_0^N D^p f(X_t)\mathrm{d}\pi_{p+1}(Z_t(b))\,.$$

The result is obtained by repeating the integration by parts.

Preliminary EML formula

Proposition (B.,Friz, Paycha '24+)

For any $b \in T_1^p(V)$

$$I_N(f,X) = \int_0^N f(Y_t)\mathrm{d}X_t - [f(X_t)b^1]_{t=0}^N$$
$$+ \sum_{l=2}^{p}(-1)^l[D^{l-1}f(X_t)\pi_l(Z_t(b))]_{t=0}^N + R^p(b)$$

where the remainder $R^m(b)$ *is given by the Stieltjes integral*

$$R^p(b) = (-1)^{p+1}\int_0^N D^p f(X_t)\mathrm{d}\pi_{p+1}(Z_t(b))\,.$$

The result is obtained by repeating the integration by parts.

Three fundamental questions

Even if the formula appears to be the same as before, three main questions need to be understood:

- How can I best describe the sawtooth signature $Z_t(b)$?
- How do you choose b to find periodic Bernoulli polynomials in a non-periodic context?
- What formula is obtained when $R^p(b) = 0$?

It is natural to compare $Z(b)$ with the signature of the path X
$S(X)\colon \{0 < s < t < N\} \to T_1((V))$

$$\pi_l(S_{s,t}(X)) = \int_{\Delta^l_{s,t}} \mathrm{d}X_{s_1} \otimes \cdots \otimes \mathrm{d}X_{s_l} \quad \Delta^l_{s,t} = \{s \leq s_1 \leq \cdots \leq s_l \leq t\}$$

The process $t \to S_{s,t}(X)$ is the solution of the differential equation

$$\begin{cases} dS_{s,t}(X) = S_{s,t}(X) \otimes \mathrm{d}X_t \\ S_{s,s}(X) = \mathbf{1} \end{cases}$$

and it is the basis for rough paths

It is natural to compare $Z(b)$ with the signature of the path X
$S(X)\colon \{0 < s < t < N\} \to T_1((V))$

$$\pi_l(S_{s,t}(X)) = \int_{\Delta^l_{s,t}} \mathrm{d}X_{s_1} \otimes \cdots \otimes \mathrm{d}X_{s_l} \quad \Delta^l_{s,t} = \{s \leq s_1 \leq \cdots \leq s_l \leq t\}$$

The process $t \to S_{s,t}(X)$ is the solution of the differential equation

$$\begin{cases} dS_{s,t}(X) = S_{s,t}(X) \otimes \mathrm{d}X_t \\ S_{s,s}(X) = \mathbf{1} \end{cases}$$

and it is the basis for rough paths

Classic signature and sawtooth signature

Theorem (B.,Friz, Paycha '24+)

For all $b \in T_1((V))$ and $l \geq 1$ we have the identity

$$\pi_l(Z_t(b)) = (-1)^l \pi_l(S_{0,t}(X)^{-1}) \otimes b - \sum_{k=0}^{[t]} (-1)^{l-1} \pi_{l-1}(S_{k,t}(X)^{-1}) \otimes \delta X_k$$

- The dependence of b is linear with respect to $\otimes$.
- $Z_t(b)$ is an algebraic functional of the signature.

Theorem (B.,Friz, Paycha '24+)

For all $b \in T_1((V))$ and $l \geq 1$ we have the identity

$$\pi_l(Z_t(b)) = (-1)^l \pi_l(S_{0,t}(X)^{-1}) \otimes b - \sum_{k=0}^{[t]} (-1)^{l-1} \pi_{l-1}(S_{k,t}(X)^{-1}) \otimes \delta X_k$$

- The dependence of b is linear with respect to $\otimes$.
- $Z_t(b)$ is an algebraic functional of the signature.

Idea of Proof

To describe $Z_t(b)$ we couple a classical induction with the study of the reversed signature $t \to S^{\flat}_{s,t}(X)$ which solves

$$\begin{cases} dS^{\flat}_{s,t}(X) = \mathrm{d}X_t \otimes S^{\flat}_{s,t}(X) \\ S_{s,s}(X) = \mathbf{1} \end{cases}$$

the components of $S^{\flat}_{s,t}(X)$ are the reversed iterated integrals

$$\pi_I(S^{\flat}_{s,t}(X)) = \int_{\Delta^{l,*}_{0,t}} dX_{s_1} \otimes \cdots \otimes dX_{s_l} \, , \quad \Delta^{l,*}_{0,t} = \{0 \leq s_l < \ldots < s_1 \leq t\}$$

We can calculate $\pi_I(S^{flat}_{s,t}(X))$ using the relation

$$\pi_I(S^{\flat}_{s,t}(X)) = (-1)^l \pi_I(S_{s,t}(X)^{-1})$$

Idea of Proof

To describe $Z_t(b)$ we couple a classical induction with the study of the reversed signature $t \to S^{\flat}_{s,t}(X)$ which solves

$$\begin{cases} dS^{\flat}_{s,t}(X) = \mathrm{d}X_t \otimes S^{\flat}_{s,t}(X) \\ S_{s,s}(X) = \mathbf{1} \end{cases}$$

the components of $S^{\flat}_{s,t}(X)$ are the reversed iterated integrals

$$\pi_I(S^{\flat}_{s,t}(X)) = \int_{\Delta^{l,*}_{0,t}} dX_{s_1} \otimes \cdots \otimes dX_{s_l} \,, \quad \Delta^{l,*}_{0,t} = \{0 \leq s_l < \ldots < s_1 \leq t\}$$

We can calculate $\pi_I(S^{flat}_{s,t}(X))$ using the relation

$$\pi_I(S^{\flat}_{s,t}(X)) = (-1)^l \pi_I(S_{s,t}(X)^{-1})$$

Optimal choice of initial constants

It is natural to look for the constant b which minimises $R^p(b)$.

By decomposing R^p symmetrically we apply Cauchy-Schwarz

$$\mathbb{E}\|R^p(b)\| \lesssim \left(\mathbb{E}\int_0^N \|\pi_p(Z_t(b))\|^2_{V^{\otimes p}} |\mathrm{d}X_t|\right)^{1/2}$$

with $|\mathrm{d}X_t|$ the total variation of the process

Optimal choice of initial constants

It is natural to look for the constant b which minimises $R^p(b)$.

By decomposing R^p symmetrically we apply Cauchy-Schwarz

$$\mathbb{E}\|R^p(b)\| \lesssim \left(\mathbb{E}\int_0^N \|\pi_p(Z_t(b))\|^2_{V^{\otimes p}} |\mathrm{d}X_t|\right)^{1/2}$$

with $|\mathrm{d}X_t|$ the total variation of the process

Optimal choice of initial constants

As we are looking for a sequence of universal constants $b \in T_1((V))$ we would like to define $\pi_p(b)$ recursively

Definition (B.,Friz, Paycha '24+)

For all $l \geq 1$ we set

$$\pi_l(b^*) = -\frac{1}{\mathbb{E}\int_0^N |\mathrm{d}X_t|}\mathbb{E}\int_0^N \pi_l(Z_t(b^{*,<l}))|\mathrm{d}X_t|$$

where $b^{*,<l} = (1, \pi_1(b^*), \ldots, \pi_{l-1}(b^*), 0, \ldots)$. We call $b^* \in T_1((V))$ the optimal average constants

Optimal choice of initial constants

As we are looking for a sequence of universal constants $b \in T_1((V))$ we would like to define $\pi_p(b)$ recursively

Definition (B.,Friz, Paycha '24+)

For all $l \geq 1$ we set

$$\pi_l(b^*) = -\frac{1}{\mathbb{E}\int_0^N |\mathrm{d}X_t|}\mathbb{E}\int_0^N \pi_l(Z_t(b^{*,<l}))|\mathrm{d}X_t|$$

where $b^{*,<l} = (1, \pi_1(b^*), \ldots, \pi_{l-1}(b^*), 0, \ldots)$. We call $b^* \in T_1((V))$ the optimal average constants

Optimal choice of initial constants

Theorem (B.,Friz, Paycha '24+)

When V is a Hilbert space the value $\pi_l(b^*)$ minimizes for all $l \geq 1$ the functional

$$v \in V^{\otimes l} \to \mathbb{E}\int_0^N \|\pi_l(Z_t(b^{*,<l} + v))\|^2_{V^{\otimes l}} |\mathrm{d}X_t| \,.$$

Thanks to the linear dependence of b, we can prove that

$$\mathbb{E}\int_0^N \|\pi_l(Z_t(b^{*,<l} + v))\|^2_{V^{\otimes l}} |\mathrm{d}X_t| = \mathbb{E}\int_0^N \|\pi_l(Z_t(b^{*,<l})) + v\|^2_{V^{\otimes l}} |\mathrm{d}X_t|$$

The result follows from the optimisation of a quadratic functional

Optimal choice of initial constants

Theorem (B.,Friz, Paycha '24+)

When V is a Hilbert space the value $\pi_l(b^*)$ minimizes for all $l \geq 1$ the functional

$$v \in V^{\otimes l} \to \mathbb{E}\int_0^N \|\pi_l(Z_t(b^{*,<l}+v))\|^2_{V^{\otimes l}}|\mathrm{d}X_t|\,.$$

Thanks to the linear dependence of b, we can prove that

$$\mathbb{E}\int_0^N \|\pi_l(Z_t(b^{*,<l}+v))\|^2_{V^{\otimes l}}|\mathrm{d}X_t|$$
$$= \mathbb{E}\int_0^N \|\pi_l(Z_t(b^{*,<l}))+v\|^2_{V^{\otimes l}}|\mathrm{d}X_t|$$

The result follows from the optimisation of a quadratic functional

Moyennes optimales et nombres de Bernoulli

Theorem (B.,Friz, Paycha '24+)

When $X_t = t$ and $V = \mathbb{R}$ for all $l \geq 2$ we have

$$\pi_l(Z_t(b^*)) = Q_l(t)$$

The proof is derived by an explicit calculation and an intrinsic property of the Bernoulli numbers

In general, this choice should improve the numerical efficiency in the computation of Riemann-Stieltjes integrals.

Theorem (B.,Friz, Paycha '24+)

When $X_t = t$ and $V = \mathbb{R}$ for all $l \geq 2$ we have

$$\pi_l(Z_t(b^*)) = Q_l(t)$$

The proof is derived by an explicit calculation and an intrinsic property of the Bernoulli numbers

In general, this choice should improve the numerical efficiency in the computation of Riemann-Stieltjes integrals.

Zero Remainder in EML

A classical application of EML are the Faulhaber formulas when $f(x) = x^q$, $q \geq 1$ et $p = q + 1$

$$\sum_{k=1}^{N-1} k^q = \frac{N^{q+1}}{q+1} + \sum_{l=1}^{q} \frac{B_l}{l!} \frac{q!}{(q-l+1)} N^{q-l+1}$$

When $\mathbf{b} = \mathbf{1}$ et $D^p f = 0$ the preliminary EML becomes

$$\int_0^N f(X_s) dX_s = \sum_{k=0}^{N-1} f(X_k) \delta X_k + \sum_{l=2}^{p} (-1)^{l+1} D^{l-1} f(X_N) \pi_l(Z_N(\mathbf{1}))$$

A classical application of EML are the Faulhaber formulas when $f(x) = x^q$, $q \geq 1$ et $p = q + 1$

$$\sum_{k=1}^{N-1} k^q = \frac{N^{q+1}}{q+1} + \sum_{l=1}^{q} \frac{B_l}{l!} \frac{q!}{(q-l+1)} N^{q-l+1}$$

When $\mathbf{b} = \mathbf{1}$ et $D^p f = 0$ the preliminary EML becomes

$$\int_0^N f(X_s) dX_s = \sum_{k=0}^{N-1} f(X_k) \delta X_k + \sum_{l=2}^{p} (-1)^{l+1} D^{l-1} f(X_N) \pi_l(Z_N(\mathbf{1}))$$

Zero Remainder in EML

When $f(X)(Y) = X \otimes Y$

$$\int_{\Delta^2_{0,N}} \mathrm{d}X_{s_1} \otimes \mathrm{d}X_{s_2} = \sum_{k=0}^{N-1} X_k \otimes \delta X_k - \pi_2(Z_N)$$
$$= \sum_{0 \leq l < k < N} \delta X_l \otimes \delta X_k - \pi_2(Z_N)$$

The term on the right involves the discrete signature of order 2 defined from the time series $\{X_k\}_{k=0,\ldots,N}$

The sawtooth signature can be seen as a correction between discrete and continuous signatures.

Zero Remainder in EML

When $f(X)(Y) = X \otimes Y$

$$\int_{\Delta^2_{0,N}} \mathrm{d}X_{s_1} \otimes \mathrm{d}X_{s_2} = \sum_{k=0}^{N-1} X_k \otimes \delta X_k - \pi_2(Z_N)$$
$$= \sum_{0 \leq l < k < N} \delta X_l \otimes \delta X_k - \pi_2(Z_N)$$

The term on the right involves the discrete signature of order 2 defined from the time series $\{X_k\}_{k=0,\ldots,N}$

The sawtooth signature can be seen as a correction between discrete and continuous signatures.

Given a time series $x\colon [\![0, N]\!] \to V$, consider the iterated sums $\Sigma(x)\colon \{0 \leq m < n \leq N\} \to T_1((V))$

$$\pi_I(\Sigma_{m,n}(x)) = \sum_{m \leq j_1 < \ldots < j_l < n} \delta x_{j_1} \otimes \ldots \otimes \delta x_{j_n}$$

Unfortunately, these tensors are not sufficient to describe all possible products.

$$\pi_1(\Sigma_{0,N}(x)) \otimes \pi_1(\Sigma_{0,N}(x)) = x_N - x_0 \otimes x_N - x_0$$

$$= \sum_{0 \leq l < k < N} \delta x_l \otimes \delta x_k + \sum_{0 \leq k < l < N} \delta x_l \otimes \delta x_k + \sum_{k=0}^{N-1} (\delta x_k)^{\otimes 2}$$

Discrete Signatures

Given a time series $x\colon [\![0, N]\!] \to V$, consider the iterated sums $\Sigma(x)\colon \{0 \leq m < n \leq N\} \to T_1((V))$

$$\pi_I(\Sigma_{m,n}(x)) = \sum_{m \leq j_1 < \ldots < j_l < n} \delta x_{j_1} \otimes \ldots \otimes \delta x_{j_n}$$

Unfortunately, these tensors are not sufficient to describe all possible products.

$$\pi_1(\Sigma_{0,N}(x)) \otimes \pi_1(\Sigma_{0,N}(x)) = x_N - x_0 \otimes x_N - x_0$$
$$= \sum_{0 \leq l < k < N} \delta x_l \otimes \delta x_k + \sum_{0 \leq k < l < N} \delta x_l \otimes \delta x_k + \sum_{k=0}^{N-1} (\delta x_k)^{\otimes 2}$$

Discrete Signatures

For any integer $n > 0$ a composition of n is a vector $I = (i_1, , \ldots , i_k)$ of positive integers s.t. $i_1 + \ldots i_k = n$.

$$\|I\| = \sum_{j=1}^{k} i_j \,, \quad |I| = k$$

Definition (DET 2020, BP2023)

We define the discrete signature as the value $\{\Sigma^I(x) \colon \{0 \leq m < n \leq N\} \to V^{\otimes \|I\|}\}_{I \in Composition}$

$$\Sigma^I_{m,n}(x) = \sum_{m \leq j_1 < \ldots < j_{|I|} < n} (\delta x_{j_1})^{\otimes i_1} \otimes \ldots (\delta x_{j_1})^{\otimes i_{|I|}}$$

For any integer $n > 0$ a composition of n is a vector $I = (i_1, , \ldots , i_k)$ of positive integers s.t. $i_1 + \ldots i_k = n$.

$$\|I\| = \sum_{j=1}^{k} i_j \,, \quad |I| = k$$

Definition (DET 2020, BP2023)

We define the *discrete signature* *as the value* $\{\Sigma^I(x) \colon \{0 \le m < n \le N\} \to V^{\otimes \|I\|}\}_{I \in Composition}$

$$\Sigma^I_{m,n}(x) = \sum_{m \le j_1 < \ldots < j_{|I|} < n} (\delta x_{j_1})^{\otimes i_1} \otimes \ldots (\delta x_{j_1})^{\otimes i_{|I|}}$$

Discrete Signatures and linear interpolations

Theorem (Chen 1954, AFS 2019, DET 2020)

Let $X^{lin}: [0, N] \to V$ be the linear interpolation of a time series $x: [\![0, N]\!] \to V$. For any integer $u \geq 1$ we have

$$\pi_u(S_{m,n}(X^{lin})) = \sum_{I \in C(u)} \frac{1}{i_1! \dots i_k!} \Sigma^I_{m,n}(x)$$

Thanks to a slight generalisation of the identity

$$\int_0^N f(X_s) dX_s = \sum_{k=0}^{N-1} f(X_k)\delta X_k + \sum_{l=2}^{p} (-1)^l D^{l-1} f(X_N) \pi_l(Z_N)$$

we can provide an alternative proof of this identity with possible extensions.

Theorem (Chen 1954, AFS 2019, DET 2020)

Let $X^{lin}\colon [0, N] \to V$ be the linear interpolation of a time series $x\colon [\![0, N]\!] \to V$. For any integer $u \geq 1$ we have

$$\pi_u(S_{m,n}(X^{lin})) = \sum_{I \in C(u)} \frac{1}{i_1! \dots i_k!} \Sigma^I_{m,n}(x)$$

Thanks to a slight generalisation of the identity

$$\int_0^N f(X_s) dX_s = \sum_{k=0}^{N-1} f(X_k)\delta X_k + \sum_{l=2}^{p} (-1)^l D^{l-1} f(X_N)\pi_l(Z_N)$$

we can provide an alternative proof of this identity with possible extensions.

Remarks and Perspectives

- Similar results can be obtained for right-point Riemann sums

$$\sum_{k=0}^{N-1} f(X_{k+1})\delta X_k$$

- Extension to semimartingales, rough paths.

- Numerical simulations of integrals?

谢谢！