PHYSICAL REVIEW D **87,** 034511 (2013)

Shadow Hamiltonians, Poisson brackets, and gauge theories

A. D. Kennedy*
School of Physics & Astronomy and SUPA, The University of Edinburgh, The King's Buildings, Edinburgh EH9 3JZ, United Kingdom

P. J. Silva†
Centro de Física Computacional, Departamento de Física, Universidade de Coimbra, Rua Larga, Coimbra 3004-516, Portugal

M. A. Clark‡
NVIDIA Corporation, 2701 San Tomas Expressway, Santa Clara, California 95050, USA
(Received 24 October 2012; published 19 February 2013)

Numerical lattice gauge theory computations to generate gauge field configurations including the effects of dynamical fermions are usually carried out using algorithms that require the molecular dynamics evolution of gauge fields using symplectic integrators. Sophisticated integrators are commonly used but hard to optimize, and force-gradient integrators show promise especially for large lattice volumes. We explain that symplectic integrators lead to very efficient Monte Carlo algorithms because they exactly conserve a shadow Hamiltonian. The shadow Hamiltonian may be expanded in terms of Poisson brackets and can be used to optimize the integrators. We show how this may be done for gauge theories by extending the formulation of Hamiltonian mechanics on Lie groups to include Poisson brackets and shadows and by giving a general method for the practical computation of forces, force gradients, and Poisson brackets for gauge theories.

DOI: 10.1103/PhysRevD.87.034511 PACS numbers: 11.15.Ha, 02.70.Ns

I. INTRODUCTION

Essentially all algorithms used in lattice gauge theory computations to generate gauge field configurations including the effects of dynamical fermions are variants of the Hybrid Monte Carlo (HMC) algorithm [1], which requires a reversible and area-preserving integrator for its molecular dynamics step. The simplest such method is the leapfrog integrator, but there is a large class of *symplectic integrators* [2] that have these properties and are potentially more cost effective. Indeed, many state-of-the-art computations use the second-order minimum norm integrator [3–5] which has a free parameter, which has been tuned heretofore in an *ad hoc* manner.

The formulation of Hamiltonian dynamics on Lie group manifolds, which is required for molecular dynamics on gauge fields [6,7], and the fact that symplectic integrators conserve a shadow Hamiltonian are well known; our goal is to combine the two and show how to construct the shadow Hamiltonian for gauge theories. This is most easily done using the formalism of differential forms [8–12]; in order to fix our notation and establish the necessary results, some of which are not easy to find in the literature, we provide a brief review in Appendix A.

The shadow Hamiltonian is expressed as an asymptotic expansion in the integration step size $\delta\tau$ whose coefficients depend on the parameters specifying the integrator under consideration and a collection of Poisson brackets. These Poisson brackets are complicated functions on phase space, where in the case of gauge field molecular dynamics a point in phase space is an entire gauge field configuration and its associated "fictitious" momenta. For extensive systems such as field theories, unlike the few body systems considered previously [13,14], the values of the Poisson brackets have a distribution that is sharply peaked about their mean values when we choose the starting points of their molecular dynamics trajectories to be chosen from the distribution e^{-H}, as is done in the HMC algorithm. This may be understood as a consequence of the central limit theorem applied to the contributions to the Poisson brackets coming from many independent regions of space-time. This means that for configurations that occur with non-vanishingly small probability, the shadow Hamiltonian may be considered to be a function of the average values of the Poisson brackets; if these are measured on a few test trajectories then the integrator parameters may be chosen to minimize the computational cost [15–17]. Perhaps surprisingly this does not correspond to minimizing the average difference between the Hamiltonian and its shadow[1] but instead to minimizing the variance of the distribution of the shadow. We shall not be concerned with the details of this tuning procedure here, but we refer the interested

*adk@ph.ed.ac.uk
†psilva@teor.fis.uc.pt
‡mclark@nvidia.com

[1]Since the shadow is only defined up to an additive constant this cannot be too surprising.

reader to Ref. [17] for details: instead, the aim of this paper is to explain how the Poisson brackets, forces, and force gradients may be computed at any given point in phase space.

In Ref. [7] expressions for the molecular dynamics force were derived from the classical mechanics specified by the Hamiltonian function and a suitable chosen group-invariant fundamental two-form. We extend this analysis to obtain an expression for the force-gradient for gauge fields [18], which can be used to provide a "second derivative" integrator step for the construction of improved integrators [13,14].

A. Multiple link updates

For much of this paper we shall be considering a Hamiltonian system with a phase space which is the cotangent bundle $T^*\mathcal{G}$ over a base space that is a Lie group manifold $\mathcal{G}$ and whose fibres are isomorphic to its Lie algebra. We shall call the cotangent one-forms "momenta," although in the context of HMC they are called "fictitious momenta" as they are quite different from the canonical momenta of the underlying field theory. For a gauge field theory we may associate such a phase space with every link of the lattice. One might at first think that we need to introduce some fibre bundle structure over the space-time lattice itself but fortunately that is not necessary. We can consider the molecular dynamics evolution of each gauge link separately; they are coupled together through the action that plays the role of the potential energy part of the Hamiltonian, but the kinetic energy part does not couple different links. For HMC we are free to choose the form of the kinetic energy, so we take it to be of the form[2] $T(p) = \frac{1}{2}\sum_\ell c_\ell p_\ell^2$ where p_ℓ is the momentum associated with the link ℓ, and c_ℓ is a link-dependent coefficient that is constant in molecular dynamics "fictitious time". If we wish to evolve the single link ϕ_ℓ on its own we choose $c_{\ell'} = \delta_{\ell,\ell'}$ so that $\dot{\phi}_{\ell'} = \partial H/\partial p_{\ell'} = \partial T/\partial p_\ell = c_{\ell'} p_{\ell'} = 0$ if $\ell \neq \ell'$. We are also free to choose $c_\ell = 1$ for all links, which is the usual situation where we update the gauge field simultaneously across the entire lattice. Another interesting choice for the kinetic energy is to choose $c_\ell = 1$ for all spatial links and $c_\ell = \xi$ for all temporal ones: this is the procedure suggested in Refs. [19,20] for evolving anisotropic lattices.[3] The momentum anisotropy ξ is a parameter that can be adjusted to optimize the HMC algorithm for a given anisotropy in the action; if the spatial and temporal contributions to the Poisson brackets are measured separately then the techniques of Clark *et al.* [17] can be used to tune ξ along with other integrator parameters.

B. Pseudofermion forces

So far we have only been discussing pure gauge theories, but in practice the cost of most lattice computations is dominated by the inclusion of fermions. This is because we need to solve a large system of linear equations in order to update the fictitious momenta (i.e., to apply the Hamiltonian vector field $\hat{\boldsymbol{S}}$ in the notation we will introduce later). Typically we have an action S which is the sum of a pure gauge part S_G, built out of sums of small Wilson loops (traces of a closed loops of gauge links) such as plaquettes and a pseudofermion part S_F built out of sums of pairs of pseudofermion fields ϕ connected by a string of gauge links. If we want to compute the force acting on a particular gauge link[4] U then it is convenient to write $S_G = \mathrm{Re}\,\mathrm{tr}(\sqcap U)$ and $S_F = \phi^\dagger \mathcal{M}^{-1}(U)\phi$ where the "staple" $\sqcap$ is the sum of all gauge link strings that connect the ends of the link U that correspond to the Wilson loops in S_G, and the Hermitian lattice matrix $\mathcal{M}(U)$ is the sum of all gauge link strings that include U that occur in S_F. For a local action all of these strings are in some neighbourhood of U, and we have dropped all other terms in the action because they are independent of U and therefore do not contribute to the force on that particular link. In reality, we update many or all the links on the lattice at once, so we compute the force on each link in parallel. By the "force" we mean the quantity $\boldsymbol{e}_i(S)T^i$ where $\boldsymbol{e}_i$ is a linear differential operator (vector field) whose action on U is specified by $\boldsymbol{e}_i(U) = -T_i U$ and which we shall define carefully later (15), and T^i is the representation of a generator of the gauge group. It is important to note that here $\boldsymbol{e}_i$ acts only on the gauge link U; it gives zero if applied to any other link variable. There is an opportunity for confusion when we refer to $\boldsymbol{e}_i$ as a vector field; it is a vector field defined over the phase space of the link U, but it is *not* a field over the space-time lattice. In order to reduce confusion we refer to quantities defined over the space-time lattice as *lattice vectors*, and space-time linear differential operators such as the Dirac operator (or more precisely lattice difference operators acting on lattice vectors such as the Wilson-Dirac operator) as *lattice matrices*.

The contribution to the force from the pure gauge part of the action is $\boldsymbol{e}_i(S_G)T^i = \mathrm{Re}\,\mathrm{tr}(\sqcap \boldsymbol{e}_i(U))T^i = -\mathrm{Re}\,\mathrm{tr}(\sqcap T_i U)T^i = -\mathrm{Re}\,\mathrm{tr}(U \sqcap T_i)T^i = -a\mathcal{T}(U\sqcap)$, $\mathcal{T}$ being the projector onto the Lie algebra, that is $\mathcal{T}(X) = \mathrm{Re}\,\mathrm{tr}(XT_i)T^i/a$ where there is an implicit sum over i as usual and the generators T_i are normalized such that $\mathrm{tr}(T_i T_j) = a\delta_{ij}$. If the gauge group is SU(N) and we choose its generators to be anti-Hermitian so as not to

[2]For notational simplicity we consider here a theory with a scalar field ϕ and the corresponding momentum p defined on the links of a lattice.

[3]In Ref. [20] the temporal step size is adjusted rather than the kinetic energy, which is equivalent after a rescaling of the temporal momenta.

[4]We shall refer both to a gauge link variable and the link on which it lives as U when there is no ambiguity.

introduce artificial factors of i, then $\mathcal{T}(X)$ is just the traceless anti-Hermitian part of X.

The pseudofermion contribution to the force is $\boldsymbol{e}_i(S_F)T^i = \phi^\dagger \boldsymbol{e}_i(\mathcal{M}(U)^{-1})\phi T^i$. Since $\boldsymbol{e}_i$ is a linear differential operator we have $0 = \boldsymbol{e}_i(\mathbb{I}) = \boldsymbol{e}_i(\mathcal{M}\mathcal{M}^{-1}) = \boldsymbol{e}_i(\mathcal{M})\mathcal{M}^{-1} + \mathcal{M}\boldsymbol{e}_i(\mathcal{M}^{-1})$, and hence $\boldsymbol{e}_i(\mathcal{M}^{-1}) = -\mathcal{M}^{-1}\boldsymbol{e}_i(\mathcal{M})\mathcal{M}^{-1}$. Therefore $\boldsymbol{e}_i(S_F)T^i = -\mathrm{Re}\,\mathrm{tr}[\boldsymbol{e}_i(\mathcal{M}(U))X \otimes X^\dagger]T^i$, where we have defined $X \equiv \mathcal{M}^{-1}\phi$ to be the solution of a large but sparse system of linear equations (since $\mathcal{M}$ is local); this may be computed on all lattice sites and used to update some or all gauge links in parallel. The outer product $X \otimes X^\dagger$ is the rank one Hermitian lattice matrix whose action on an arbitrary lattice vector y is proportional to the projection of y along X, namely $(X \otimes X^\dagger)y = X(X^\dagger y)$.

We can express the pseudofermion action in the form $S_F = -\mathrm{Re}\,\mathrm{tr}[\mathcal{M}(U)X \otimes X^\dagger]$ analogous to that of $S_\mathcal{G}$ if we consider X to be a lattice vector that is independent of U. This means that once we have computed X the calculation of the gauge and pseudofermion parts of the force and related quantities are very similar. Both the gauge and pseudofermion actions can be written as the trace of lattice operators times U, where the lattice operators are either local ($\sqcap$ and $\mathcal{M}$) or low rank ($X \otimes X^\dagger$). Both local and low rank operators are relatively cheap to apply to lattice vectors or to trace, the former only involving links in the neighbourhood of U, and the latter only involving inner products of lattice vectors. For example, we may evaluate the trace $\mathrm{tr}[\mathcal{M}(U)X \otimes X^\dagger] = X^\dagger\mathcal{M}(U)X$ as the inner product of $X^\dagger$ with the vector $\mathcal{M}(U)X$.

If we include spin degrees of freedom then we must replace $X \otimes X^\dagger$ by a sum of outer products for each spin component, but the result is still a low rank matrix which is therefore cheap to apply. Likewise, if we wish to introduce n pseudofermion fields so as to reduce the noise in the stochastic estimate of the fermionic force and thus defer the breakdown in the asymptotic expansion for the shadow Hamiltonian to significantly larger integrator step sizes [21–25] then we only increase the rank by a factor of n.

C. Outline

The structure of this paper is as follows. In Sec. II we consider the general formulation of Hamiltonian mechanics on a symplectic manifold [26]; this serves to introduce the important concepts of the fundamental two-form, the Hamiltonian vector field it associates with any zero-form, and the Poisson bracket of two zero-forms. We show that Poisson brackets satisfy the Jacobi identity and that the commutator of two Hamiltonian vector fields is itself a Hamiltonian vector field; we explain the isomorphism between the Lie algebra of commutators of Hamiltonian vector fields and that of Poisson brackets of zero-forms. The reason we need all this mathematical machinery is that when we consider Hamiltonian mechanics on Lie groups in Sec. IV we will introduce a nontrivial fundamental two-form in order to make the dynamics symmetric under the action of the group. Moreover, the fact that Hamiltonian vector fields form a Lie algebra is crucial for the definition of the shadow Hamiltonian, which we give in Sec. III. The exposition assumes some knowledge of the theory of differential forms, an overview of which is given in Appendix A.

Section III introduces symplectic integrators by noting that if a zero-form on phase space only depends on the momenta p or only on the positions q then the integral curves of its Hamiltonian vector field are easily found. We are interested in Hamiltonians $H(q, p) = T(p) + S(q)$ that are the sum of two such functions, and we show how this allows us to construct symplectic integrators to find approximate integral curves for $\hat{\boldsymbol{H}}$ using the Baker-Campbell-Hausdorff (BCH) formula. We give some simple examples of integrators for a system on a symplectic manifold with fundamental two-form $\boldsymbol{\omega} = \boldsymbol{d}q \wedge \boldsymbol{d}p$ and show how to compute the corresponding shadow Hamiltonians. When the kinetic energy is of the form $T(p) = \frac{1}{2}p^2$ we show that the Poisson bracket $\{S, \{S, T\}\}$ is independent of p and explain how it may thus be used to construct a force-gradient integrator step.

Section IV defines a symplectic structure on Lie group manifolds, or more precisely on their cotangent bundle $T^*\mathcal{G}$, that is compatible with the group structure. This is done by introducing the natural fundamental two-form terms of Maurer-Cartan forms, and it is here that the mathematical framework we have developed becomes necessary. We derive explicit formulas for Hamiltonian vector fields and Poisson brackets in terms of the momentum coordinates (which are well defined globally) and the family of left-invariant vector fields dual to the Maurer-Cartan forms. All the independent Poisson brackets of S and T that can occur in shadow Hamiltonians up to and including $\mathcal{O}(\delta\tau^4)$ are given explicitly for the case where S is momentum independent and T is quadratic in the momenta. We then show how to express the results in terms of matrix representations of the Lie group, as these are what are used in practice.

In Sec. V we evaluate the formulas for the Poisson brackets for the physically interesting case of the fundamental representation of SU(N). We show that they can all be expressed as traces of a collection of Lie-algebra-valued quantities: as these live on links we name them basic lattice vectors.

In Sec. VI we address the problem of computing these basic lattice vectors. We do this first for the simple case where only a single link is updated and then introduce the algebra of *towers* to give an efficient way of computing them in general.

Appendix A gives a brief survey of the theory of differential forms and serves to fix our notation and conventions, as does Appendix B which gives an overview of the properties of Lie groups.

II. HAMILTONIAN MECHANICS

A. Symplectic manifolds

A Hamiltonian system is defined on *phase space*, which is a differential manifold $\mathcal{M}$ with a *symplectic structure* given by some *fundamental two-form* $\boldsymbol{\omega}$ that is closed, $\boldsymbol{d\omega} = 0$, and globally invertible. Phase space is usually the cotangent bundle $T^*\mathcal{G}$ over some *configuration space* manifold $\mathcal{G}$. For every zero-form $F \in \Lambda^0$ on $\mathcal{M}$, that is for every C^∞ smooth function $F\colon \mathcal{M} \to \mathcal{M}$, there is a corresponding *Hamiltonian vector field* $\hat{\boldsymbol{F}} \in \mathrm{Ham}\,\mathcal{M}$ such that $\boldsymbol{dF} \equiv i_{\hat{\boldsymbol{F}}}\boldsymbol{\omega}$: in other words, $\boldsymbol{dF}(\boldsymbol{y}) = (i_{\hat{\boldsymbol{F}}}\boldsymbol{\omega})(\boldsymbol{y}) = \boldsymbol{\omega}(\hat{\boldsymbol{F}}, \boldsymbol{y})$ for any vector field $\boldsymbol{y}$.

A zero-form Z corresponds to a vanishing Hamiltonian vector field if and only if $\boldsymbol{dZ} = 0$. We have the following short exact sequence $0 \to \mathbb{R} \to \Lambda^0(\mathcal{M}) \to \mathrm{Ham}\,\mathcal{M} \to 0$, which implies that there is a bijective diffeomorphism $\Lambda^0(\mathcal{M})/\mathbb{R} \leftrightarrow \mathrm{Ham}\,\mathcal{M}$. The nature of this correspondence between zero-forms (up to an additive constant) and Hamiltonian vector fields will be examined further in the following sections.

B. Poisson brackets

Consider the action of a Hamiltonian vector field $\hat{\boldsymbol{F}}$ on a zero-form G,

$$\hat{\boldsymbol{F}}G = \boldsymbol{dG}(\hat{\boldsymbol{F}}) = i_{\hat{\boldsymbol{G}}}\boldsymbol{\omega}(\hat{\boldsymbol{F}}) = \boldsymbol{\omega}(\hat{\boldsymbol{G}}, \hat{\boldsymbol{F}}) \equiv \{F, G\},$$

where in the first equality we have made use of the definition of the exterior derivative of a zero-form G acting on an arbitrary vector field $\boldsymbol{y}$, $\boldsymbol{dG}(\boldsymbol{y}) \equiv \boldsymbol{y}G$, and in the last equality we have introduced the *Poisson bracket* $\{A, B\} \equiv -\boldsymbol{\omega}(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}})$ for any pair of zero-forms A and B. The minus sign has to appear somewhere, and our convention is to introduce it here in the definition of the Poisson bracket.

C. Jacobi identity

The invariant expression (A3) for the exterior derivative $\boldsymbol{d\omega}$ of a two-form $\boldsymbol{\omega}$ applied to three arbitrary vector fields $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$

$$\begin{aligned}\boldsymbol{d\omega}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = {} & \boldsymbol{x}\boldsymbol{\omega}(\boldsymbol{y}, \boldsymbol{z}) + \boldsymbol{y}\boldsymbol{\omega}(\boldsymbol{z}, \boldsymbol{x}) + \boldsymbol{z}\boldsymbol{\omega}(\boldsymbol{x}, \boldsymbol{y}) \\ & - \boldsymbol{\omega}([\boldsymbol{x}, \boldsymbol{y}], \boldsymbol{z}) - \boldsymbol{\omega}([\boldsymbol{y}, \boldsymbol{z}], \boldsymbol{x}) \\ & - \boldsymbol{\omega}([\boldsymbol{z}, \boldsymbol{x}], \boldsymbol{y})\end{aligned}$$

displays an interesting cyclic symmetry in the three vector fields $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$. This has an important consequence if $\boldsymbol{\omega}$ is the fundamental two-form and the vector fields are Hamiltonian: if A, B, and C are three arbitrary zero-forms then

$$\hat{\boldsymbol{A}}\boldsymbol{\omega}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}) = -\hat{\boldsymbol{A}}\{B, C\} = -\{A, \{B, C\}\},$$

and also

$$\begin{aligned}\boldsymbol{\omega}([\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}], \hat{\boldsymbol{C}}) &= -\boldsymbol{\omega}(\hat{\boldsymbol{C}}, [\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}]) = -\boldsymbol{dC}([\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}]) \\ &= -[\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}]C = (\hat{\boldsymbol{B}}\,\hat{\boldsymbol{A}} - \hat{\boldsymbol{A}}\,\hat{\boldsymbol{B}})C \\ &= \{B, \{A, C\}\} - \{A, \{B, C\}\}.\end{aligned}$$

We thus find that the condition $\boldsymbol{d\omega} = 0$ implies that the cyclic sum of nested Poisson brackets must vanish, $\boldsymbol{d\omega}(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}) = \{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0$: this is just the *Jacobi identity* which, together with the antisymmetry of the Poisson bracket demonstrates that zero-forms on $\mathcal{M}$ together with the product given by the Poisson bracket form a *Lie algebra*.

We can use the Jacobi identity to derive another useful result. The commutator of any two vector fields is a vector field [Eq. (A2)]; if both vector fields are Hamiltonian then their commutator is also Hamiltonian since

$$\begin{aligned}[\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}]C &= (\hat{\boldsymbol{A}}\,\hat{\boldsymbol{B}} - \hat{\boldsymbol{B}}\,\hat{\boldsymbol{A}})C = \{A, \{B, C\}\} - \{B, \{A, C\}\} \\ &= -\{C, \{A, B\}\} = \{\{A, B\}, C\} = \widehat{\{\boldsymbol{A}, \boldsymbol{B}\}}C,\end{aligned}$$

where we applied the Jacobi identity in the antepenultimate step. Since this must hold $\forall\ C \in \Lambda^0$, we have

$$[\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}] = \widehat{\{\boldsymbol{A}, \boldsymbol{B}\}} \in \mathrm{Ham}\,\mathcal{M} \tag{1}$$

telling us that not only is the commutator of two Hamiltonian vector fields Hamiltonian as promised, but also that it corresponds to the zero-form that is the Poisson bracket of the zero-forms corresponding to the original pair of Hamiltonian vector fields. The bijection $\Lambda^0(\mathcal{M})/\mathbb{R} \leftrightarrow \mathrm{Ham}(\mathcal{M})$ is therefore an isomorphism of Lie algebras.

D. Lie derivatives and equations of motion

Given a Hamiltonian $H \in \Lambda^0(\mathcal{M})$ and a fundamental two-form $\boldsymbol{\omega}$ we may construct the Hamiltonian vector field $\hat{\boldsymbol{H}}$, and for any point $p \in \mathcal{M}$ we may, at least locally, define an *integral curve*. We may also define a *local flow* $\sigma\colon \mathcal{I} \times \mathcal{U} \to \mathcal{M}$ of trajectories starting at any point $p \in \mathcal{U} \subseteq \mathcal{M}$ in some neighbourhood of p, $\sigma\colon \mathbb{R} \to \mathcal{M}$, satisfying Hamilton's equations $d\sigma/dt = \hat{\boldsymbol{H}}$ and the initial condition $\sigma(0) = p$. Hamilton's equations are thus most naturally expressed in terms of Lie derivatives (Appendix A 5), $d\boldsymbol{T}/dt = \mathcal{L}_{\hat{\boldsymbol{H}}}\boldsymbol{T}$, for any tensor $\boldsymbol{T}$. In particular, a scalar field (zero-form) F, vector field $\boldsymbol{v}$, and one-form $\boldsymbol{\theta}$ must obey

$$\begin{aligned}\frac{dF}{dt} &= \mathcal{L}_{\hat{\boldsymbol{H}}}F = \hat{\boldsymbol{H}}F = \{H, F\}, \\ \frac{d\boldsymbol{v}}{dt} &= \mathcal{L}_{\hat{\boldsymbol{H}}}\boldsymbol{v} = [\hat{\boldsymbol{H}}, \boldsymbol{v}], \quad \text{and} \\ \frac{d\boldsymbol{\theta}}{dt} &= \mathcal{L}_{\hat{\boldsymbol{H}}}\boldsymbol{\theta} = (i_{\hat{\boldsymbol{H}}}\boldsymbol{d} + \boldsymbol{d}i_{\hat{\boldsymbol{H}}})\boldsymbol{\theta}.\end{aligned}$$

The formal solution of the equation of motion $d\boldsymbol{T}/dt = \mathcal{L}_{\hat{\boldsymbol{H}}}\boldsymbol{T}$ is $\boldsymbol{T}(t) = \exp(t\mathcal{L}_{\hat{\boldsymbol{H}}})\boldsymbol{T}(0)$, where the

exponential function is defined as $\exp(t\mathcal{L}_{\hat{\boldsymbol{H}}}) = \lim_{n\to\infty}(1 + \frac{t}{n}\mathcal{L}_{\hat{\boldsymbol{H}}})^n = \sum_{j=0}^{\infty}(t\mathcal{L}_{\hat{\boldsymbol{H}}})^j/j!$.

III. SYMPLECTIC INTEGRATORS AND SHADOW HAMILTONIANS

A. Baker-Campbell-Hausdorff formula

The BCH formula states that if A and B belong to an associative algebra then

$$\ln(e^A e^B) = \sum_{n=1}^{\infty} c_n(A, B), \tag{2}$$

where the c_n, belonging to the *free Lie algebra*,[5] are recursively determined from the relations $c_1 = A + B$ and

$$(n+1)c_{n+1} = \sum_{m=1}^{\lfloor n/2\rfloor} \frac{B_{2m}}{(2m)!} \sum_{\substack{k_1,\ldots,k_{2m}\geq 1 \\ k_1+\cdots+k_{2m}=n}} \operatorname{ad} c_{k_1} \ldots \operatorname{ad} c_{k_{2m}}(A+B) - \frac{1}{2}(\operatorname{ad} c_n)(A-B) \quad \text{for } n \geq 1, \tag{3}$$

where $\operatorname{ad} a\colon b \mapsto [a, b]$ and the *Bernoulli numbers* B_n are defined by

$$\frac{x}{e^x - 1} \equiv \sum_{n\geq 0} \frac{B_n x^n}{n!}.$$

The first few terms in the Hausdorff series are

$$\ln(e^A e^B) = (A+B) + \frac{1}{2}[A,B] + \frac{1}{12}([A,[A,B]] - [B,[A,B]]) - \frac{1}{24}[B,[A,[A,B]]] + \frac{1}{720}\begin{pmatrix} -4[B,[A,[A,[A,B]]]] & -6[[A,B],[A,[A,B]]] \\ +4[B,[B,[A,[A,B]]]] & -2[[A,B],[B,[A,B]]] \\ -[A,[A,[A,[A,B]]]] & +[B,[B,[B,[A,B]]]] \end{pmatrix} + \cdots$$

From this we easily obtain the corresponding formula for a symmetric product

$$\ln(e^{A/2}e^B e^{A/2}) = (A+B) - \frac{1}{24}(2[B,[A,B]] + [A,[A,B]]) + \frac{1}{5760}\begin{pmatrix} 32[B,[B,[A,[A,B]]]] & -16[[A,B],[B,[A,B]]] \\ +28[B,[A,[A,[A,B]]]] & +12[[A,B],[A,[A,B]]] \\ +8[B,[B,[B,[A,B]]]] & +7[A,[A,[A,[A,B]]]] \end{pmatrix} + \cdots \tag{4}$$

B. Symplectic integrators

The integral curve of a Hamiltonian vector field $\hat{\boldsymbol{A}}$ is given by the exponential map $t \mapsto \exp(t\hat{\boldsymbol{A}})$ acting on the initial point. Given two Hamiltonian vector fields $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ we can construct a curve that is alternately tangential to each vector field from the composition of their exponential maps $t \mapsto [\exp(t\hat{\boldsymbol{A}}/n)\exp(t\hat{\boldsymbol{B}}/n)]^n$ for some $n \in \mathbb{N}$. Such a map is called a *symplectic integrator* as it manifestly preserves the symplectic structure because it is a composition of exponential maps each of which preserves this structure. The BCH formula (2) tells us that this curve itself is in fact the integral curve of a vector field $\boldsymbol{D}_{t/n}$

$$\begin{aligned}\left[\exp\left(\frac{t\hat{\boldsymbol{A}}}{n}\right)\exp\left(\frac{t\hat{\boldsymbol{B}}}{n}\right)\right]^n &= \left[\exp\left((\hat{\boldsymbol{A}}+\hat{\boldsymbol{B}})\frac{t}{n} + \sum_{m=2}^{\infty} c_m(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})\left(\frac{t}{n}\right)^m\right)\right]^n \\ &= \exp\left[\left(\hat{\boldsymbol{A}}+\hat{\boldsymbol{B}} + \sum_{m=2}^{\infty} c_m(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})\left(\frac{t}{n}\right)^{m-1}\right)t\right] \\ &= \exp(\boldsymbol{D}_{t/n}t),\end{aligned}$$

where $\boldsymbol{D}_\varepsilon \equiv \hat{\boldsymbol{A}} + \hat{\boldsymbol{B}} + \sum_{m=2}^{\infty} c_m(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})\varepsilon^{m-1}$. As all the c_m are commutators, Eq. (1) tells us that $\boldsymbol{D}_\varepsilon$ is a Hamiltonian vector field corresponding to the *shadow* zero-form D_ε under the isomorphism $\operatorname{Ham}\mathcal{M} \leftrightarrow \Lambda_0(\mathcal{M})/\mathbb{R}$ discussed earlier. In other words, $\boldsymbol{D}_\varepsilon = \hat{\boldsymbol{D}}_\varepsilon$, where the zero-form $D_\varepsilon \equiv A + B + \sum_{m=2}^{\infty} c'_m(A,B)\varepsilon^{m-1}$ with the c'_m defined by (3) in terms of the Poisson bracket image of the adjoint under the Lie algebra isomorphism (1) $\operatorname{ad}\hat{\boldsymbol{A}} \mapsto \widehat{\operatorname{ad}}A$, where $\widehat{\operatorname{ad}}A\colon B \mapsto \{A, B\}$. We note in passing that the shadow is only defined up to an additive constant.

The BCH formula is obtained by formal manipulation of the exponential series, so we should choose a sufficiently large n to ensure that the Hausdorff series converges. In order to study the convergence of the BCH formula we need to specify a topology on the space of Hamiltonian vector fields $\operatorname{Ham}\mathcal{M}$. It is simpler to ask the same question about the convergence of the corresponding expansion for the shadow Hamiltonian, for which there is an obvious topology as the coefficients are zero-forms and we can use the usual L_p norms. In most cases none of these norms are bounded so the series is only asymptotic at best. In HMC the momenta are selected from a Gaussian distribution $e^{-T(p)}$ so the values of the Poisson brackets can become arbitrarily large but with exponentially small probability. There is no value of ε for which the Hausdorff series always converges, but it might well be that for any

[5] That is the Lie algebra whose Lie bracket is the commutator constructed from the associative product. For more details about free Lie algebras and a proof of the BCH formula see Appendix B of Ref. [27].

$\delta > 0$ we can find an $\varepsilon > 0$ such that it does converge with probability $> 1 - \delta$. This may be acceptable for HMC, where an exponentially small chance of a trajectory becoming unstable is unimportant: it will presumably be rejected and the next momentum or pseudofermion refreshment will resolve the problem. If the large norm comes from the gauge field configuration then there could be more severe problems.

C. Symmetric symplectic integrators

In general a symplectic integrator is not reversible, that is the group commutator

$$\exp(-t\hat{\boldsymbol{A}}/n)\exp(-t\hat{\boldsymbol{B}}/n)\exp(t\hat{\boldsymbol{A}}/n)\exp(t\hat{\boldsymbol{B}}/n) \neq \mathbb{I};$$

indeed we immediately see from this expression that that the integrator is reversible if and only if $[\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}] = 0$. This blemish is easily eradicated by using a *symmetric symplectic integrator* such as $\exp(t\hat{\boldsymbol{A}}/2n)\exp(t\hat{\boldsymbol{B}}/n) \times \exp(t\hat{\boldsymbol{A}}/2n)$. An additional advantage of such integrators is that only even powers of ε occur in the Hausdorff series for their shadow Hamiltonians D_ε, so $A + B - D_\varepsilon = \mathcal{O}(\varepsilon^2)$, making them better approximations to the exponential map of $\hat{\boldsymbol{A}} + \hat{\boldsymbol{B}}$ itself.

D. Practical integrators

Finding a closed-form expression for the integral curve of some Hamiltonian vector field $\hat{\boldsymbol{A}}$ is impossible in most cases as there is no closed-form solution of Hamilton's equations. However, there are some special cases where we can find such a solution.

For example, suppose that in some local patch of phase space with coordinates q and p the fundamental two-form is[6] $\boldsymbol{\omega} = \boldsymbol{d}q \wedge \boldsymbol{d}p$, A is an arbitary zero-form, and $\boldsymbol{X}$ is an arbitrary vector field on phase space. Then

$$\boldsymbol{d}A = \frac{\partial A}{\partial q}\boldsymbol{d}q + \frac{\partial A}{\partial p}\boldsymbol{d}p, \qquad \boldsymbol{X} \equiv X_q \frac{\partial}{\partial q} + X_p \frac{\partial}{\partial p},$$
$$\hat{\boldsymbol{A}} \equiv A_q \frac{\partial}{\partial q} + A_p \frac{\partial}{\partial p},$$

and we have

$$\begin{aligned}\boldsymbol{d}A(\boldsymbol{X}) &= \frac{\partial A}{\partial q}X_q + \frac{\partial A}{\partial p}X_p = \boldsymbol{\omega}(\hat{\boldsymbol{A}}, \boldsymbol{X}) \\ &= (\boldsymbol{d}q \wedge \boldsymbol{d}p)\left(A_q \frac{\partial}{\partial q} + A_p \frac{\partial}{\partial p}, X_q \frac{\partial}{\partial q} + X_p \frac{\partial}{\partial p}\right) \\ &= A_q X_p - A_p X_q.\end{aligned}$$

Since $\boldsymbol{X}$ is arbitrary we can equate coefficients of X_q and X_p to obtain

$$\hat{\boldsymbol{A}} = \frac{\partial A}{\partial p}\frac{\partial}{\partial q} - \frac{\partial A}{\partial q}\frac{\partial}{\partial p}.$$

Let $c(t) = (q_t, p_t)$ be the integral curve of $\hat{\boldsymbol{A}}$ with $c(0) = (q_0, p_0)$, which means that for any zero-form f it must satisfy the differential equations

$$(\hat{\boldsymbol{A}}f) \circ c = \frac{d}{dt}(f \circ c),$$

or equivalently

$$\dot{q}_t = \frac{\partial A}{\partial p}(q_t, p_t) \quad \text{and} \quad \dot{p}_t = -\frac{\partial A}{\partial q}(q_t, p_t),$$

which are Hamilton's equations if A is the Hamiltonian.

Now, suppose that $A(q, p) = T(p)$ is only a function of the momenta, then $\hat{\boldsymbol{T}} = T'(p)\partial/\partial q$, and Hamilton's equations reduce to the pair $\dot{q}_t = T'(p)$ and $\dot{p}_t = 0$ of first-order differential equations with constant coefficients, with the solution that the momentum is constant, $p_t = p_0$, and $q_t = q_0 + T'(p_0)t$ grows linearly in t. The case where $A(q, p) = S(q)$ is analogous. If we have a function $A(q, p) = H(q, p) = T(p) + S(q)$, perhaps the Hamiltonian itself, that can be decomposed into the sum of a kinetic energy and a potential energy then we can easily integrate either term separately, and we can use a symplectic integrator to approximate the integral curves of $\hat{\boldsymbol{H}}$ itself.

In fact we have established a stronger result, namely we can find the exact integral curves of a shadow Hamiltonian H_ε that differs from H by terms of $\mathcal{O}(\varepsilon)$ in closed form. A symplectic integrator thus not only exactly preserves the symplectic structure but also conserves the value of H (the energy) up to order ε for arbitrarily long times: unfortunately, the integral curves of $\hat{\boldsymbol{H}}$ and $\hat{\boldsymbol{H}}_\varepsilon$ usually diverge from each other after a relatively short time despite this. This happens even if their equations of motion are not chaotic: symplectic integrators are very good at conserving energy and phase space volume, but they are not particularly good in finding the correct trajectory through phase space.

For HMC applications where we only care about exact reversibility, exact area preservation, and good energy conservation we see that symmetric symplectic integrators meet all the requirements, and the divergence of the shadow integral curves from the true ones is unimportant.

Given the fundamental two-form $\boldsymbol{\omega} = \boldsymbol{d}q \wedge \boldsymbol{d}p$ we may evaluate the Poisson bracket of two arbitrary zero-forms A and B, namely

$$\begin{aligned}\{A, B\} &\equiv -\boldsymbol{\omega}(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}) \\ &= -(\boldsymbol{d}q \wedge \boldsymbol{d}p)\left(\frac{\partial A}{\partial p}\frac{\partial}{\partial q} - \frac{\partial A}{\partial q}\frac{\partial}{\partial p}, \frac{\partial B}{\partial p}\frac{\partial}{\partial q} - \frac{\partial B}{\partial q}\frac{\partial}{\partial p}\right) \\ &= \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} - \frac{\partial A}{\partial q}\frac{\partial B}{\partial p}.\end{aligned}$$

[6] We can always find coordinates for which this is true according to Darboux's theorem.

For the Hamiltonian $H(q, p) = T(p) + S(q)$ any integrator constructed from $e^{\varepsilon\hat{\boldsymbol{S}}}$ and $e^{\varepsilon\hat{\boldsymbol{T}}}$ steps will conserve a shadow whose BCH expansion may be expressed in terms of the Poisson brackets

$$\{S, T\} = -S'T' \qquad \{S, \{S, T\}\} = -S'\frac{\partial\{S, T\}}{\partial p} = S'^2T''$$

$$\{T, \{S, T\}\} = T'\frac{\partial\{S, T\}}{\partial q} = -S''T'^2,$$

and so forth.

For example, the leapfrog integrator $[\exp(\frac{1}{2}\delta\tau\hat{\boldsymbol{S}}) \times \exp(\delta\tau\hat{\boldsymbol{T}})\exp(\frac{1}{2}\delta\tau\hat{\boldsymbol{S}})]^{t/\delta\tau}$ is the simplest symmetric symplectic integrator (there is a variant in which $\hat{\boldsymbol{S}}$ and $\hat{\boldsymbol{T}}$ are interchanged). From (4) we find that it conserves the shadow Hamiltonian

$$\begin{aligned}\tilde{H} &= T + S - \frac{\delta\tau^2}{24}(\{S, \{S, T\}\} + 2\{T, \{S, T\}\}) + \mathcal{O}(\delta\tau^4)\\ &= H - \frac{\delta\tau^2}{24}(S'^2T'' - 2S''T'^2) + \mathcal{O}(\delta\tau^4).\end{aligned}$$

E. Higher-order integrators

Let us briefly give some simple examples of more complicated integrators. The second-order minimum norm integrator [3–5] is

$$\left[\exp(\lambda\delta\tau\hat{\boldsymbol{S}})\exp\left(\frac{1}{2}\delta\tau\hat{\boldsymbol{T}}\right)\exp((1-2\lambda)\delta\tau\hat{\boldsymbol{S}}) \times \exp\left(\frac{1}{2}\delta\tau\hat{\boldsymbol{T}}\right)\exp(\lambda\delta\tau\hat{\boldsymbol{S}})\right]^{t/\delta\tau}$$

with shadow

$$\tilde{H} = T + S + \delta\tau^2\left(\frac{6\lambda^2 - 6\lambda + 1}{12}\{S, \{S, T\}\} + \frac{1-6\lambda}{24}\{T, \{S, T\}\}\right) + \mathcal{O}(\delta\tau^4),$$

and it has the free parameter λ as well as the integration step size $\delta\tau$.

It is interesting to note that if as is usual the kinetic energy is quadratic, $T(p) = \frac{1}{2}p^2$, then the Poisson bracket $\{S, \{S, T\}\} = S'^2$ is independent of the momentum p, and thus we can find the integral curve of its Hamiltonian vector field $\{\boldsymbol{S}, \widehat{\{\boldsymbol{S}, \boldsymbol{T}\}}\} = -2S'S''\partial/\partial p$. The corresponding integrator step $e^{\varepsilon\widehat{\{S,\{S,T\}\}}}$ is called a *force-gradient* integrator step, because it involves second derivatives of the potential S.

We can use the force-gradient step to define a force-gradient integrator

$$\left[\exp\left(\frac{1}{6}\delta\tau\hat{\boldsymbol{S}}\right)\exp\left(\frac{1}{2}\delta\tau\hat{\boldsymbol{T}}\right) \times \exp\left(\frac{1}{72}[48\delta\tau\hat{\boldsymbol{S}} - \delta\tau^3\{\boldsymbol{S}, \widehat{\{\boldsymbol{S}, \boldsymbol{T}\}}\}]\right) \times \exp\left(\frac{1}{2}\delta\tau\hat{\boldsymbol{T}}\right)\exp\left(\frac{1}{6}\delta\tau\hat{\boldsymbol{S}}\right)\right]^{t/\delta\tau}$$

with shadow

$$\tilde{H} = T + S - \frac{\delta\tau^4}{155520}\begin{pmatrix} 41\{S, \{S, \{S, \{S, T\}\}\}\} \\ +36\{\{S, T\}, \{S, \{S, T\}\}\} \\ +72\{\{S, T\}, \{T, \{S, T\}\}\} \\ +84\{T, \{S, \{S, \{S, T\}\}\}\} \\ +126\{T, \{T, \{S, \{S, T\}\}\}\} \\ +54\{T, \{T, \{T, \{S, T\}\}\}\} \end{pmatrix} + \mathcal{O}(\delta\tau^6), \tag{5}$$

where we have chosen the integrator parameters to eliminate all terms of $\mathcal{O}(\delta\tau^2)$ in the shadow. The Poisson bracket $\{S, \{S, \{S, T\}\}\} = 0$ so the first and fourth Poisson brackets in (5) are also identically zero, however, formula (5) is valid more generally. Note that the middle step has combined the Hamiltonian vector fields $\hat{\boldsymbol{S}}$ and $\{\boldsymbol{S}, \widehat{\{\boldsymbol{S}, \boldsymbol{T}\}}\}$ because they commute.

There is no compelling reason to choose the parameters to eliminate the $\delta\tau^2$ errors: in general we should introduce some parameters constrained only by the conditions that the leading-order term in the shadow should be the original Hamiltonian and that the total step size should be $\delta\tau$, and then adjust these parameters to minimize the cost of our integrator for the specific problem it is being applied to. On the other hand, we can build integrators whose leading error is $\delta\tau^4$ (or $\delta\tau^{2n}$ for any n for that matter), without requiring force-gradient steps. Nevertheless, integrators with force-gradient steps may be cheaper than those without: it would be surprising if the optimal coefficient of the force-gradient term was exactly zero.

In HMC for lattice field theory H and $\tilde{H}$ are extensive quantities, that is they are proportional to the lattice volume V for sufficiently large V, so the leading error is proportional to $V\delta\tau^{2n}$ if $H - \tilde{H} = \mathcal{O}(\delta\tau^{2n})$. In order to keep the Monte Carlo acceptance rate fixed we therefore need to vary $\delta\tau \propto V^{-1/2n}$, and as the cost $Vt/\delta\tau$ of a trajectory of length t is proportional to the number of steps and the volume, we may estimate that the cost varies as $V^{1+1/2n}$. Of course there are many other contributions to the cost that have been ignored, but for large enough V this suggests that we want to increase n.

IV. HAMILTONIAN MECHANICS ON LIE GROUPS

A. Fundamental two-form on a Lie group

The cotangent bundle $T^*\mathcal{G}$ over any manifold $\mathcal{G}$ has a natural symplectic structure. For the case where $\mathcal{G}$ is a Lie

group a point in $T^*\mathcal{G}$ may be written as $(g, \boldsymbol{p})$, where $g \in \mathcal{G}$ and $\boldsymbol{p} \in T^*\mathcal{G}(g)$ is called the *momentum* or *Liouville* form. As explained in Appendix B, the vectors in tangent space at the identity $T\mathcal{G}(\mathbb{I})$ correspond to the Lie algebra of left-invariant vector fields $\boldsymbol{e}_i$ on $\mathcal{G}$, and their dual one-forms $\boldsymbol{\theta}^i$ satisfy the Maurer-Cartan equations. The momentum may be written in the Maurer-Cartan basis as $\boldsymbol{p} = p_i\boldsymbol{\theta}^i$, where $\boldsymbol{p}(\boldsymbol{e}_j) = p_i\boldsymbol{\theta}^i(\boldsymbol{e}_j) = p_i\delta^i_j = p_j$. We shall choose the fundamental two-form to be

$$\boldsymbol{\omega} \equiv -\boldsymbol{d}\boldsymbol{p} = -\boldsymbol{d}(p_i\boldsymbol{\theta}^i), \tag{6}$$

and using the Maurer-Cartan equations it may be written as

$$\boldsymbol{\omega} = \boldsymbol{\theta}^i \wedge dp_i + \frac{1}{2}p_i c^i_{jk}\boldsymbol{\theta}^j \wedge \boldsymbol{\theta}^k.$$

If F is a zero-form on the cotangent bundle $T^*\mathcal{G}$ then the corresponding Hamiltonian vector field $\hat{\boldsymbol{F}} = F^i\boldsymbol{e}_i + \bar{F}_i\partial/\partial p_i$ in $TT^*\mathcal{G}$ is defined by $\boldsymbol{d}F = i_{\hat{\boldsymbol{F}}}\boldsymbol{\omega}$, or $\boldsymbol{d}F(\boldsymbol{y}) = \boldsymbol{\omega}(\hat{\boldsymbol{F}}, \boldsymbol{y})$ for all vector fields $y = y^i\boldsymbol{e}_i + \bar{y}_i\partial/\partial p_i$. Expanding this expression gives

$$\begin{aligned}\boldsymbol{d}F(\boldsymbol{y}) = \boldsymbol{y}F &= \boldsymbol{e}_i(F)y^i + \frac{\partial F}{\partial p_i}\bar{y}_i = \boldsymbol{\omega}(\hat{\boldsymbol{F}}, \boldsymbol{y}) \\ &= F^i\bar{y}_i - y^i\bar{F}_i + p_i c^i_{jk}F^j y^k,\end{aligned}$$

so equating the coefficients of y^i and $\bar{y}_i$ we find $\boldsymbol{e}_i(F) = -\bar{F}_i + p_j c^j_{ki}F^k$ and $\partial F/\partial p_i = F^i$. We thus find that the vector field $\hat{\boldsymbol{F}}$ is

$$\hat{\boldsymbol{F}} = \frac{\partial F}{\partial p_i}\boldsymbol{e}_i + \left(p_j c^j_{ki}\frac{\partial F}{\partial p_k} - \boldsymbol{e}_i(F)\right)\frac{\partial}{\partial p_i}. \tag{7}$$

From this we can evaluate the Poisson bracket of two arbitrary Hamiltonian vector fields corresponding to zero-forms F and G,

$$\begin{aligned}\{F, G\} &\equiv -\boldsymbol{\omega}(\hat{\boldsymbol{F}}, \hat{\boldsymbol{G}}) \\ &= p_i c^i_{jk}\frac{\partial F}{\partial p_j}\frac{\partial G}{\partial p_k} + \frac{\partial F}{\partial p_i}\boldsymbol{e}_i(G) - \frac{\partial G}{\partial p_i}\boldsymbol{e}_i(F).\end{aligned} \tag{8}$$

B. Hamiltonian vector fields for T and S

For HMC we may take the Hamiltonian to be of the form $H = T + S$ where the kinetic energy T: $T^*\mathcal{G} \to \mathbb{R}$ is a function only of the momenta which we may choose to be of the form

$$T = \frac{1}{2}\langle \boldsymbol{p}, \boldsymbol{p}\rangle = \frac{1}{2}p^i p_i \tag{9}$$

using the Cartan-Killing metric (Appendix B 5). Hence $\partial T/\partial p_i = p^i$, and the potential energy S: $\mathcal{G} \to \mathbb{R}$ is a function only of the group parameters.

For the kinetic and potential energy zero-forms the corresponding vector fields are thus

$$\hat{\boldsymbol{T}} = p^i\boldsymbol{e}_i + c^j_{ki}p_j p^k\frac{\partial}{\partial p_i} = p^i\boldsymbol{e}_i \quad \text{and} \quad \hat{\boldsymbol{S}} = -\boldsymbol{e}_i(S)\frac{\partial}{\partial p_i} \tag{10}$$

using (7), where we have made use of the total antisymmetry of the structure constants for a semisimple Lie algebra, $c^j_{ki}p_j p^k = c_{jki}p^j p^k = 0$.

C. Poisson brackets of S and T

We may compute the Poisson brackets of S and T from (8)

$$\{S, T\} = -p^i\boldsymbol{e}_i(S) \tag{11}$$

$$\{S,\{S,T\}\} = \boldsymbol{e}^i(S)\boldsymbol{e}_i(S) \quad \{T,\{S,T\}\} = -p^i p^j\boldsymbol{e}_i\boldsymbol{e}_j(S) \tag{12}$$

$$\begin{aligned}\{T,\{S,\{S,T\}\}\} &= 2p^i\boldsymbol{e}_i\boldsymbol{e}_j(S)\boldsymbol{e}^j(S) \\ \{S,\{S,\{S,T\}\}\} &= 0 \\ \{T,\{T,\{S,T\}\}\} &= -p^i p^j p^k\boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k(S) \\ \{T,\{T,\{S,\{S,T\}\}\}\} &= 2p^i p^j\boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k(S)\boldsymbol{e}^k(S) \\ &\quad + 2p^i p^j\boldsymbol{e}_i\boldsymbol{e}_k(S)\boldsymbol{e}_j\boldsymbol{e}^k(S) \\ \{\{S,T\},\{T,\{S,T\}\}\} &= c^k_{ij}p^i p^\ell\boldsymbol{e}^j(S)[\boldsymbol{e}_k\boldsymbol{e}_\ell(S) + \boldsymbol{e}_\ell\boldsymbol{e}_k(S)] \\ &\quad + p^i p^j[\boldsymbol{e}^k(S)\boldsymbol{e}_k\boldsymbol{e}_i\boldsymbol{e}_j(S) \\ &\quad - \boldsymbol{e}^k\boldsymbol{e}_i(S)\boldsymbol{e}_k\boldsymbol{e}_j(S) - \boldsymbol{e}_i\boldsymbol{e}^k(S)\boldsymbol{e}_k\boldsymbol{e}_j(S)] \\ \{T,\{S,\{S,\{S,T\}\}\}\} &= 0 \\ \{\{S,T\},\{S,\{S,T\}\}\} &= -2\boldsymbol{e}^i(S)\boldsymbol{e}^j(S)\boldsymbol{e}_i\boldsymbol{e}_j(S) \\ \{T,\{T,\{T,\{S,T\}\}\}\} &= -p^i p^j p^k p^\ell\boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k\boldsymbol{e}_\ell(S) \\ \{S,\{S,\{S,\{S,T\}\}\}\} &= 0.\end{aligned} \tag{13}$$

Observe that according to Eq. (12) $\{S,\{S,T\}\}$ does not depend on the momentum, so just as in Sec. III E we can use it to define a force-gradient integrator step corresponding to the Hamiltonian vector field

$$\{\boldsymbol{S}, \widehat{\{\boldsymbol{S}, \boldsymbol{T}\}}\} = -\boldsymbol{e}_i(\boldsymbol{e}^j(S)\boldsymbol{e}_j(S))\frac{\partial}{\partial p_i}. \tag{14}$$

D. Representations

If U: $\mathcal{G} \to \mathrm{Gl}(n, \mathbb{C}) \equiv \mathrm{Aut}\mathbb{C}^N$ is a matrix representation of $\mathcal{G}$ then it satisfies $U(gh) = U(g)U(h)$ for all $g, h \in \mathcal{G}$. We may view any matrix element U_{ab} of the representation as a complex valued zero-form as it is well defined over the entire group manifold. The left action L_g: $h \mapsto gh$ induces the map L_{g*}: $U_{ab} \mapsto U_{ab} \circ L_g$ according to the definition given in Appendix A 4, so $(L_{g*}U_{ab})(h) = U_{ab}(gh) = [U(g)U(h)]_{ab} = \sum_{c=1}^n U_{ac}(g)U_{cb}(h)$ for all h, or equivalently $L_{g*}U_{ab} = \sum_{c=1}^n U_{ac}(g)U_{cb}$. In other words, the map L_{g*} takes the zero-form U_{ab} to a linear combination of zero-forms U_{cb} with coefficients $U_{ac}(g) \in \mathbb{C}$. We can express this more succinctly by considering U to be a matrix-valued zero-form, whence $L_{g*}U = U(g)U$.

Application of the vector field $\boldsymbol{e}_i$ to U gives a matrix-valued zero-form $\boldsymbol{e}_i U$ whose value at some point $g \in \mathcal{G}$ is $\boldsymbol{e}_i U(g) = L_{g*}\boldsymbol{e}_i U(\mathbb{I})$. $\boldsymbol{e}_i$ is left-invariant $L_g^*\boldsymbol{e}_i = \boldsymbol{e}_i$, so we have $L_{g*}\boldsymbol{e}_i U = L_{g*}L_g^*\boldsymbol{e}_i U = L_{g*}L_{g^{-1}*}\boldsymbol{e}_i L_{g*}U = \boldsymbol{e}_i L_{g*}U = \boldsymbol{e}_i U(g)U = U(g)\boldsymbol{e}_i U$. This allows us to evaluate $\boldsymbol{e}_i U$ at any point g in terms of the value of $\boldsymbol{e}_i U$ at the identity. Defining the *generators* of the representation as $T_i \equiv \boldsymbol{e}_i U(\mathbb{I})$, we obtain $\boldsymbol{e}_i U(g) = U(g)\boldsymbol{e}_i U(\mathbb{I}) = U(g)T_i$ or more succinctly $\boldsymbol{e}_i U = UT_i$.

As on the one hand $[\boldsymbol{e}_i, \boldsymbol{e}_j]U = c_{ij}^k \boldsymbol{e}_k U = c_{ij}^k UT_k$, and on the other $[\boldsymbol{e}_i, \boldsymbol{e}_j]U = \boldsymbol{e}_i\boldsymbol{e}_j U - \boldsymbol{e}_j\boldsymbol{e}_i U = \boldsymbol{e}_i UT_j - \boldsymbol{e}_j UT_i = UT_iT_j - UT_jT_i = U[T_i, T_j]$, we see that the generators must satisfy the commutation relations $[T_i, T_j] = c_{ij}^k T_k$ upon multiplying on the left by U^{-1}.

Unfortunately the usual convention [7,28] is that the derivative of a link variable is

$$\boldsymbol{e}_i U = -T_i U, \tag{15}$$

and this is used in most computer implementations. This arises from considering right-invariant vector fields. Briefly, the right action on a group is defined by R_g: $h \mapsto hg$, and the induced maps by $R_{g*}U = U \circ R_g$ and $R_g^*\boldsymbol{e}_i = R_{g^{-1}*}\boldsymbol{e}_i R_{g*}$. If we assume that $\boldsymbol{e}_i$ is *right-invariant* then it satisfies $R_g^*\boldsymbol{e}_i = \boldsymbol{e}_i$, and following an argument completely analogous to that in the text we find $\boldsymbol{e}_i U(g) = R_{g*}\boldsymbol{e}_i U(\mathbb{I})$ since $g = R_g\mathbb{I} = L_g\mathbb{I}$ and $R_{g*}\boldsymbol{e}_i U = \boldsymbol{e}_i UU(g)$. We then have to define the generators by $\boldsymbol{e}_i U(\mathbb{I}) = -T_i$, leading to $\boldsymbol{e}_i U = -T_i U$. We must include the minus sign in the definition of the generators for right-invariant vector fields satisfying $[\boldsymbol{e}_i, \boldsymbol{e}_j] = c_{ij}^k \boldsymbol{e}_k$; otherwise, they would not satisfy the commutation relations $[T_1, T_j] = c_{ij}^k T_k$. In fact, the usual convention erroneously omits the minus sign, but as the commutation relations are used to derive the Maurer-Cartan equations, and thus our fundamental two-form, the sign is significant when computing high-order Poisson brackets.

E. Equations of motion

The equations of motion are most naturally expressed in terms of Lie derivatives (Appendix A 5). The Lie derivative $\mathcal{L}_{\boldsymbol{v}}\boldsymbol{T}$ of a tensor field $\boldsymbol{T}$ is its derivative along the integral curves of the vector field $\boldsymbol{v}$, and the definition of the Lie derivative given in (A4)–(A6) implicitly provides the differential equations defining these integral curves. If $\boldsymbol{v} = \hat{\boldsymbol{H}}$ is the Hamiltonian vector field for the Hamiltonian function then these are just Hamilton's equations, and we will write $\dot{\boldsymbol{T}} \equiv \mathcal{L}_{\hat{\boldsymbol{H}}}\boldsymbol{T}$.

For the case of matrix representations we consider the matrix elements to be zero-forms as we did in Sec. IV D, so we may use equation (A4) to obtain $\dot{U} \equiv \mathcal{L}_{\hat{\boldsymbol{H}}}U = \hat{\boldsymbol{H}}U$ and $\dot{P} \equiv \mathcal{L}_{\hat{\boldsymbol{H}}}P = \hat{\boldsymbol{H}}P$ where U is a matrix representation of an element of $\mathcal{G}$ and $P \equiv p^i T_i$ the corresponding matrix representation of the momentum in the Lie algebra. Taking

$$\hat{\boldsymbol{H}} = \hat{\boldsymbol{T}} + \hat{\boldsymbol{S}} = p^i\boldsymbol{e}_i - \boldsymbol{e}_i(S)\frac{\partial}{\partial p_i}$$

with the explicit forms from (10), and using the relation $\boldsymbol{e}_i(U) = -T_i U$ of (15), we find

$$\dot{U} = \hat{\boldsymbol{T}}U = p^i\boldsymbol{e}_i(U) = -p^i T_i U = -PU$$

$$\dot{P} = \hat{\boldsymbol{S}}P = -\boldsymbol{e}_i(S)\frac{\partial P}{\partial p_i} = -\boldsymbol{e}_i(S)T^i = -F_1,$$

where we have introduced the quantity $F_1 \equiv \boldsymbol{e}_i(S)T^i$ [q.v., Eq. (17)]. The solution of these equations for separate U and P updates (i.e., for a symplectic integrator) are

$$U(t) = \exp(-Pt)U(0) \quad \text{and} \quad P(t) = P(0) - tF_1.$$

The equations of motion for the force-gradient Hamiltonian vector field of (14) is

$$\begin{aligned}\dot{P} &= \{\boldsymbol{S}, \widehat{\{\boldsymbol{S}, \boldsymbol{T}\}}\}P = -\boldsymbol{e}_i(\boldsymbol{e}_j(S)\boldsymbol{e}^j(S))\frac{\partial P}{\partial p_i} \\ &= -2\boldsymbol{e}_i\boldsymbol{e}_j(S)\boldsymbol{e}^j(S)T^i = -G\end{aligned} \tag{16}$$

with $G \equiv \boldsymbol{e}_i\boldsymbol{e}_j(S)\boldsymbol{e}^i(S)T^j$ [q.v., Eq. (17)], since $[\boldsymbol{e}_i, \boldsymbol{e}_j](S)\boldsymbol{e}^j(S) = c_{kij}\boldsymbol{e}^k(S)\boldsymbol{e}^j(S) = 0$.

V. POISSON BRACKETS IN $SU(N)$

In order to compute the Poisson brackets it is useful to express them in terms of the following set of matrices that are in the representation of the Lie algebra

$$\begin{aligned}P &\equiv p_i T^i \\ F_1 &\equiv \boldsymbol{e}_i(S)T^i \\ F_2 \equiv \mathcal{P}F_1 &= p^j\boldsymbol{e}_j\boldsymbol{e}_i(S)T^i \\ F_3 \equiv \mathcal{P}^2F_1 &= p^k\boldsymbol{e}_k p^j\boldsymbol{e}_j\boldsymbol{e}_i(S)T^i \\ F_4 \equiv \mathcal{P}^3F_1 &= p^\ell\boldsymbol{e}_\ell p^k\boldsymbol{e}_k p^j\boldsymbol{e}_j\boldsymbol{e}_i(S)T^i \\ G \equiv \mathcal{F}_1F_1 &= \boldsymbol{e}^j(S)\boldsymbol{e}_j\boldsymbol{e}_i(S)T^i,\end{aligned} \tag{17}$$

$$\begin{aligned}p_i &= \mathrm{tr}(PT_i)/a \\ \boldsymbol{e}_i(S) &= \mathrm{tr}(F_1T_i)/a \\ p^j\boldsymbol{e}_j\boldsymbol{e}_i(S) &= \mathrm{tr}(F_2T_i)/a \\ p^k\boldsymbol{e}_k p^j\boldsymbol{e}_j\boldsymbol{e}_i(S) &= \mathrm{tr}(F_3T_i)/a \\ p^\ell\boldsymbol{e}_\ell p^k\boldsymbol{e}_k p^j\boldsymbol{e}_j\boldsymbol{e}_i(S) &= \mathrm{tr}(F_4T_i)/a \\ \boldsymbol{e}^j(S)\boldsymbol{e}_j\boldsymbol{e}_i(S) &= \mathrm{tr}(GT_i)/a,\end{aligned}$$

where $\mathcal{P} = p^i\boldsymbol{e}_i$ and $\mathcal{F}_1 = \boldsymbol{e}^i(S)\boldsymbol{e}_i$ are vector fields (linear differential operators) corresponding to the matrices P and F_1, respectively. For a lattice field theory $P, F_i, G, \ldots$ will also be lattice vectors, so we shall call these quantities *basic lattice vectors*.

To derive more explicit expressions for the desired Poisson brackets it is useful to use the following identities that hold for the fundamental representation of the $su(N)$

Lie algebra,[7] for arbitrary $N \times N$ matrices X, Y, Z, and $T_i \in \mathrm{su}(N)$

$$c^i_{jk}\,\mathrm{tr}(XT^j)\mathrm{tr}(YT^k) = a\,\mathrm{tr}([X,Y]T^i); \tag{18}$$

$$\mathrm{tr}(XT_i)\mathrm{tr}(YT^i) = a\left[\mathrm{tr}(XY) - \frac{1}{N}\,\mathrm{tr}X\,\mathrm{tr}Y\right]; \tag{19}$$

$$\mathrm{tr}[X,Y] = \mathrm{tr}(XY - YX) = 0; \tag{20}$$

and

$$\begin{aligned}\mathrm{tr}([X,Y]Z) &= \mathrm{tr}(XYZ - YXZ) = \mathrm{tr}(XYZ - XZY)\\ &= \mathrm{tr}(X[Y,Z]) = \mathrm{tr}([Y,Z]X) = \mathrm{tr}([Z,X]Y);\end{aligned} \tag{21}$$

from which it follows that

$$\mathrm{tr}([X,Y]X) = \mathrm{tr}([X,X]Y) = 0 \tag{22}$$

and

$$c_{ijk}\,\mathrm{tr}(XT^i)\mathrm{tr}(YT^j)\mathrm{tr}(ZT^k) = a^2\,\mathrm{tr}([X,Y]Z). \tag{23}$$

Using (23) and (21) we easily see that

$$\begin{aligned}c^k_{ij}p^i p^\ell \boldsymbol{e}^j(S)\boldsymbol{e}_\ell\boldsymbol{e}_k(S) &= \frac{1}{a^3}c_{ijk}\,\mathrm{tr}(PT^i)\mathrm{tr}(F_1T^j)\mathrm{tr}(F_2T^k)\\ &= \frac{1}{a}\,\mathrm{tr}([F_1,F_2]P),\end{aligned} \tag{24}$$

and as (18) leads to

$$\begin{aligned}p^\ell[\boldsymbol{e}_k,\boldsymbol{e}_\ell](S) &= p^\ell c^i_{k\ell}\boldsymbol{e}_i(S) = \frac{1}{a^2}c_{k\ell i}\,\mathrm{tr}(PT^\ell)\mathrm{tr}(F_1T^i)\\ &= \frac{1}{a}\,\mathrm{tr}([P,F_1]T_k)\end{aligned} \tag{25}$$

we find using (23) that

$$\begin{aligned}&c^k_{ij}p^i p^\ell \boldsymbol{e}^j(S)[\boldsymbol{e}_k,\boldsymbol{e}_\ell](S)\\ &\quad= \frac{1}{a^3}c_{ijk}\,\mathrm{tr}(PT^i)\mathrm{tr}(F_1T^j)\mathrm{tr}([P,F_1]T^k)\\ &\quad= \frac{1}{a}\,\mathrm{tr}([F_1,P]^2).\end{aligned} \tag{26}$$

Combining Eqs. (24) and (26) we obtain

$$\begin{aligned}&c^k_{ij}p^i p^\ell \boldsymbol{e}^j(S)\{\boldsymbol{e}_k\boldsymbol{e}_\ell(S) + \boldsymbol{e}_\ell\boldsymbol{e}_k(S)\}\\ &\quad= \frac{1}{a}\,\mathrm{tr}(2[F_1,F_2]P + [F_1,P]^2).\end{aligned} \tag{27}$$

We may also deduce from (25) that

$$p^i\boldsymbol{e}_k\boldsymbol{e}_i(S) = \frac{1}{a}\,\mathrm{tr}((F_2 - [F_1,P])T_k),$$

and hence

$$p^i p^j \boldsymbol{e}^k\boldsymbol{e}_i(S)\boldsymbol{e}_k\boldsymbol{e}_j(S) = \frac{1}{a}\,\mathrm{tr}((F_2 - [F_1,P])^2) \tag{28}$$

and

$$p^i p^j \boldsymbol{e}_i\boldsymbol{e}^k(S)\boldsymbol{e}_k\boldsymbol{e}_j(S) = \frac{1}{a}\,\mathrm{tr}(F_2^2 - F_2[F_1,P]). \tag{29}$$

From the identity

$$\begin{aligned}\boldsymbol{e}_k\boldsymbol{e}_i\boldsymbol{e}_j &= [\boldsymbol{e}_k,\boldsymbol{e}_i]\boldsymbol{e}_j + \boldsymbol{e}_i[\boldsymbol{e}_k,\boldsymbol{e}_j] + \boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k\\ &= c^\ell_{ki}[\boldsymbol{e}_\ell,\boldsymbol{e}_j] + c^\ell_{ki}\boldsymbol{e}_j\boldsymbol{e}_\ell + \boldsymbol{e}_i c^\ell_{kj}\boldsymbol{e}_\ell + \boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k\\ &= c^\ell_{ki}c^m_{\ell j}\boldsymbol{e}_m + c^\ell_{ki}\boldsymbol{e}_j\boldsymbol{e}_\ell + c^\ell_{kj}\boldsymbol{e}_i\boldsymbol{e}_\ell + \boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k\end{aligned}$$

we deduce that

$$\begin{aligned}&p^i p^j \boldsymbol{e}^k(S)\boldsymbol{e}_k\boldsymbol{e}_i\boldsymbol{e}_j(S)\\ &\quad= c^\ell_{ki}\boldsymbol{e}^k(S)p^i c^m_{\ell j}p^j\boldsymbol{e}_m(S) + 2c^\ell_{ki}p^i p^j\boldsymbol{e}_j\boldsymbol{e}_\ell(S)\boldsymbol{e}^k(S)\\ &\qquad+ p^i p^j\boldsymbol{e}_i\boldsymbol{e}_j\boldsymbol{e}_k(S)\boldsymbol{e}^k(S)\\ &\quad= \frac{1}{a^4}c^\ell_{ki}\,\mathrm{tr}(F_1T^k)\mathrm{tr}(PT^i)c_{\ell jm}\mathrm{tr}(PT^j)\mathrm{tr}(F_1T^m)\\ &\qquad+ \frac{2}{a^3}c_{ki\ell}\,\mathrm{tr}(PT^i)\mathrm{tr}(F_2T^\ell)\mathrm{tr}(F_1T^k)\\ &\qquad+ \frac{1}{a^2}\,\mathrm{tr}(F_3T_k)\mathrm{tr}(F_1T^k)\\ &\quad= -\frac{1}{a}\,\mathrm{tr}([F_1,P]^2 + 2[F_1,F_2]P - F_1F_3).\end{aligned} \tag{30}$$

We thus obtain the following expressions for the desired Poisson brackets

$$\begin{aligned}\{S,T\} &= -\mathrm{tr}(F_1P)/a\\ \{S,\{S,T\}\} &= \mathrm{tr}(F_1^2)/a\\ \{T,\{S,T\}\} &= -\mathrm{tr}(F_2P)/a\\ \{T,\{S,\{S,T\}\}\} &= 2\,\mathrm{tr}(F_1F_2)/a\\ \{S,\{S,\{S,T\}\}\} &= 0\\ \{T,\{T,\{S,T\}\}\} &= -\mathrm{tr}(F_3P)/a\\ \{T,\{T,\{S,\{S,T\}\}\}\} &= 2\{\mathrm{tr}(F_1F_3) + \mathrm{tr}(F_2^2)\}/a\\ \{\{S,T\},\{T,\{S,T\}\}\} &= -\mathrm{tr}(3[F_1,F_2]P\\ &\qquad+ [F_1,P]^2 - F_1F_3 + 2F_2^2)/a\end{aligned}$$

using (27)–(30)

[7] We choose to normalize the traceless anti-Hermitian generators T_i of the fundamental representation by $\mathrm{tr}(T_iT_j) = a\delta_{ij}$, where a is an arbitrary (negative) constant. For su(3) the Hermitian Gell-Mann matrices λ_i satisfy $\mathrm{tr}(\lambda_i\lambda_j) = 2\delta_{ij}$, so our choice corresponds to $T_i = \sqrt{-a/2}\,i\lambda_i$. Moreover, our definition of the kinetic energy is $T = \frac{1}{2}p_ip^i = \mathrm{tr}(P^2)/2a$, and as we observed in the introduction changing this normalization corresponds to a scaling of molecular dynamics time. One must be careful to take all these factors into account when comparing computations using different conventions.

$$\{T,\{S,\{S,\{S,T\}\}\}\} = 0$$
$$\{\{S,T\},\{S,\{S,T\}\}\} = -2\,\mathrm{tr}(F_1 G_1)/a$$
$$\{T,\{T,\{T,\{S,T\}\}\}\} = -\mathrm{tr}(F_4 P)/a$$
$$\{S,\{S,\{S,\{S,T\}\}\}\} = 0.$$

VI. BASIC LATTICE VECTORS AND TOWERS

A. Single link updates

We now consider how to evaluate the basic lattice vectors of (17). This is particularly simple to do in the case where there is only a single link variable U, or on a lattice if we choose to only update a single link by setting the coefficient of the kinetic energy to zero everywhere else, as described in Sec. I A. In this case the potential is of the form[8] $S = \mathrm{Re}\,\mathrm{tr}(UX)$ where X is some constant $N \times N$ matrix, which in general is neither in the group nor its algebra. On a lattice where we are only updating a single link X is constructed out of products of other link variables, which are themselves constant in molecular dynamics time. We find $F_1 = \boldsymbol{e}_i(S)T^i = \mathrm{Re}\,\mathrm{tr}(\boldsymbol{e}_i(U)X)T^i = -\mathrm{Re}\,\mathrm{tr}(T_i UX)T^i = -\mathrm{Re}\,\mathrm{tr}(UXT_i)T^i = -a\mathcal{T}(UX)$ where $\mathcal{T}$ projects onto the Lie algebra, i.e., the traceless anti-Hermitian part for su(N). Likewise, $F_2 = \mathcal{P}F_1 = p^j \boldsymbol{e}_j \boldsymbol{e}_i(S)T^i = \mathrm{Re}\,\mathrm{tr}(p^j \boldsymbol{e}_j \boldsymbol{e}_i(U)X)T^i = \mathrm{Re}\,\mathrm{tr}(p^j \boldsymbol{e}_j(-T_i U)X)\,T^i = -\mathrm{Re}\,\mathrm{tr}(T_i p^j \boldsymbol{e}_j(U)X)T^i = \mathrm{Re}\,\mathrm{tr}(T_i p^j T_j UX)T^i = \mathrm{Re}\,\mathrm{tr}(PUXT_i)T^i = a\mathcal{T}(PUX)$, and so forth for the remaining quantities in (17)

$$F_1 = -\mathrm{Re}\,\mathrm{tr}(UXT_i)T^i = -a\mathcal{T}(UX),$$
$$F_2 = \mathcal{P}F_1 = \mathrm{Re}\,\mathrm{tr}(PUXT_i)T^i = a\mathcal{T}(PUX),$$
$$F_3 = \mathcal{P}^2 F_1 = -\mathrm{Re}\,\mathrm{tr}(P^2 UXT_i)T^i = -a\mathcal{T}(P^2 UX),$$
$$F_4 = \mathcal{P}^3 F_1 = \mathrm{Re}\,\mathrm{tr}(P^3 UXT_i)T^i = a\mathcal{T}(P^3 UX),$$
$$G = \mathcal{F}_1 F_1 = \mathrm{Re}\,\mathrm{tr}(F_1 UXT_i)T^i = a\mathcal{T}(F_1 UX).$$

B. Lattice updates

When we have many links we trivially generalize the definition of the fundamental two-form (6) to become sums over all links

$$\boldsymbol{\omega} = -\sum_\ell \boldsymbol{d}\boldsymbol{p}(\ell) = -\sum_\ell \boldsymbol{d}(p_i(\ell)\boldsymbol{\theta}^i(\ell))$$
$$= \sum_\ell \left(\boldsymbol{\theta}^i(\ell) \wedge dp_i(\ell) + \frac{1}{2} p_i(\ell) c^i_{jk} \boldsymbol{\theta}^j(\ell) \wedge \boldsymbol{\theta}^k(\ell)\right).$$

[8]We consider the case where the action is linear in U without loss of generality because if it occurs multiple times we can transform it into a form linear in its tensor product, which can be reduced into a sum of irreducible representations. For example, the action $S = \mathrm{Re}\,\mathrm{tr}(UXUX') = \mathrm{Re}\,\mathrm{tr}[(U \otimes U)X'']$, where $(U \otimes U')_{ij,k\ell} = U_{ik}U_{j\ell}$ and $X''_{k\ell,ij} = X_{kj}X'_{\ell i}$ are $N^2 \times N^2$ matrices, and $U \otimes U$ can be reduced into a sum of two irreducible representations acting on vectors of dimensions $\frac{1}{2}N(N-1)$ and $\frac{1}{2}N(N+1)$.

We can compress the notation by letting indices such as i also range over all links: that is $i \to (i, \ell_i)$ and the implicit sum over the basis of the Lie algebra $\sum_i$ becomes an implicit double sum $\sum_{\ell_i}\sum_i$. Of course, we also need to augment the structure constants $c^k_{ij} \to c^{(k,\ell_k)}_{(i,\ell_i)(j,\ell_j)} \equiv c^k_{ij}\delta^{\ell_k}_{\ell_i}\delta^{\ell_k}_{\ell_j}$ since the Maurer-Cartan equations do not mix links. Similarly, the kinetic energy (9) becomes

$$T = \frac{1}{2}\sum_\ell c(\ell)\langle \boldsymbol{p}(\ell), \boldsymbol{p}(\ell)\rangle = \frac{1}{2}\sum_\ell c(\ell) g_{ij} p^i(\ell) p^j(\ell)$$
$$= \frac{1}{2}\sum_\ell c(\ell) p_i(\ell) p^i(\ell),$$

where, as discussed in Appendix A 1, it is convenient to introduce a separate coefficient $c(\ell)$ in the kinetic energy for each link. We can extend our compressed notation by implicitly associating a factor of $c(\ell)$ with each occurrence of the augmented Cartan-Killing metric, $g_{ij} \to g_{(i,\ell_i)(j,\ell_j)} \equiv c(\ell_i) g_{ij}\delta_{\ell_i\ell_j}$ and hence with every contracted index i. With these conventions the definition looks like (6) and (9) again. The sums propagate to the Poisson brackets where the implicit sums over the indices in Eqs. (11)–(13) and also become sums over all links, although second derivatives such as $\boldsymbol{e}_i\boldsymbol{e}_j(S)$ have bounded support for an ultralocal action. It is important to note that the implicit factor of c_{ℓ_i} associated with contracted indices means that even though $\{S,\{S,T\}\}$ does not depend on any momentum it still has a factor of $c(\ell)$ associated with each term. If we set $c(\ell') = \delta_{\ell\ell'}$ then only link ℓ will appear in Eqs. (12) and (16), and the force-gradient integrator will therefore only act on that link.

C. Towers

The situation would seem to be much more difficult when we want to update all of the link variables simultaneously; derivatives like $\boldsymbol{e}_{i_1}\ldots\boldsymbol{e}_{i_k}(S)$ depend on k links and it might appear that it will be prohibitively expensive to compute them. Fortunately, we can avoid this combinatorial explosion; the key observation is that all the Poisson brackets and forces only depend on the basic lattice vectors, and these have only a single free lattice index. To make use of this we introduce towers of basic lattice vectors: a *tower* $T(A, B)$ is a an array of basic lattice vectors $T(A, B)_i = \mathcal{A}^i B$, where A is a basic lattice vector, $\mathcal{A}$ is the vector field associated with it, B is a sum of products of gauge links, and the index $i \in \{0, \ldots, n-1\}$ where we call n the *height* of the tower.

The basic lattice vectors in (17) may be constructed from the two towers $T(P, B)$ and $T(F_1, B)$ of heights four and two, where B is the stencil of the action S. The *stencil* is the collection of all paths in the action that start with a given link. For example, in the case of lattice gauge theory without dynamical fermions the action is a sum of Wilson loops, each Wilson loop being the trace of the product of

gauge links around a closed loop. This means we can write the action as $S = \mathrm{Re\,tr}(U_l \sqcap) + S_0$ where the *staple* $\sqcap$ is the sum of products of gauge links along paths connecting the end of the link ℓ to its beginning, and S_0 is independent of U_ℓ, as in Sec. I B. The stencil in this case is $U_\ell \sqcap$. This is familiar from the computation of the force acting on U_ℓ

$$\begin{aligned} F_1(\ell) &= \boldsymbol{e}_i(S)T^i = \boldsymbol{e}_i(\mathrm{Re\,tr}U_\ell\sqcap)T^i \;= \mathrm{Re\,tr}(e_i(U_\ell)\sqcap)T^i \\ &= \mathrm{Re\,tr}(-T_iU_\ell\sqcap)T^i = -\mathrm{Re\,tr}(U_\ell \sqcap T_i)T^i \\ &= -a\mathcal{T}(U_\ell\sqcap). \end{aligned} \tag{31}$$

The thing to notice here is that we are computing the force on the gauge link U_ℓ so the index i is really the pair (i, ℓ), and thus $\boldsymbol{e}_i(U_{\ell'}) = 0$ for any other link $\ell' \neq \ell$: in particular, $\boldsymbol{e}_i(\sqcap) = 0$, $\boldsymbol{e}_i(S_0) = 0$, and $\boldsymbol{e}_i(U_\ell\sqcap) = \boldsymbol{e}_i(U_\ell)\sqcap$. Naturally, we want to compute the force acting on every gauge link, and so the stencil computation of (31) must be carried out separately for each link: these computations can be done in parallel if desired.

In order to compute the basic lattice vector $\mathcal{A}^jF_1 = \mathcal{A}^j\boldsymbol{e}_i(S)T^i$ we proceed as follows:

$$\begin{aligned} \mathcal{A}^jF_1(\ell) &= \mathcal{A}^j\boldsymbol{e}_i(S)T^i = \mathcal{A}^j\,\mathrm{Re\,tr}(-T_iU_\ell\sqcap)T^i \\ &= -\mathrm{Re\,tr}(T_i\mathcal{A}^j(U_\ell\sqcap))T^i \\ &= -\mathrm{Re\,tr}(T_iT(A, U_\ell\sqcap)_j)T^i \\ &= -a\mathcal{T}(T(A, U_\ell\sqcap)_j). \end{aligned}$$

This is easy to do if we can compute the tower $T(A, U_\ell\sqcap)$ on the stencil $U_\ell \sqcap$.

D. Algebra of towers

It is simple to construct the tower $T(A, B)$ when B is a single gauge link U; we have $T(A, U)_j = \mathcal{A}^jU = (-A)^jU$. This follows from the definitions $T(A,U)_0 = U$ and $\mathcal{A} = a^i\boldsymbol{e}_i$ where $A = a^iT_i$, so by induction $T(A, U)_{j+1} = \mathcal{A}^{j+1}U = \mathcal{A}(\mathcal{A}^jU) = \mathcal{A}(-A)^jU = a^i\boldsymbol{e}_i((-A)^jU) = (-A)^ja^i\boldsymbol{e}_i(U) = (-A)^ja^i(-T_iU) = (-A)^{j+1}U$. Indeed, this corresponds to a convenient recursive way of constructing the tower, $T(A, U)_{j+1} = (-A)T(A, U)_j$.

If B is the product[9] of two stencils $B_1 \cdot B_2$ then we may use the Leibniz rule for the derivation $\mathcal{A}$, $\mathcal{A}(B_1 \cdot B_2) = \mathcal{A}B_1 \cdot B_2 + B_1 \cdot \mathcal{A}B_2$, or more generally

$$\mathcal{A}^j(B_1 \cdot B_2) = \sum_{k=0}^{j}\binom{j}{k}\mathcal{A}^kB_1 \cdot \mathcal{A}^{j-k}B_2.$$

The tower on the product $B_1 \cdot B_2$ is thus the product of the tower on B_1 with that on B_2, $T(A, B_1 \cdot B_2) = T(A, B_1) \cdot T(A, B_2)$, where the product is defined by[10]

$$(T(A, B_1) \cdot T(A, B_2))_j = \sum_{k=0}^{j}\binom{j}{k}T(A, B_1)_k \cdot T(A, B_2)_{j-k}.$$

The tower on the sum of two stencils $B_1 + B_2$ is even simpler, since $\mathcal{A}(B_1 + B_2) = \mathcal{A}B_1 + \mathcal{A}B_2$. We just have $T(A, B_1 + B_2) = T(A, B_1) + T(A, B_2)$ where $(T(A, B_1) + T(A, B_2))_j = T(A, B_1)_j + T(A, B_2)_j$.

E. Pseudofermion towers

The principal advantage of updating all links simultaneously is when we include the effects of (pseudo) fermions in the dynamics. As described in Sec. I B this entails solving a large linear system to obtain the quantity $X = \mathcal{M}^{-1}\phi$ needed to compute the force ($\mathcal{M}$ being a lattice Dirac operator) and it is worthwhile to reuse this solution to update many links.

We therefore need to compute towers for stencils that include outer products such as $X \otimes X^\dagger$. This may be done by computing the tower $T(A, X)$ on $X = \mathcal{M}^{-1}\phi$. Observe that $\mathcal{A}\phi = 0$ as the pseudofermion lattice (site) vector ϕ does not depend on U—we want to follow the molecular dynamics evolution of the gauge links and momenta in the presence of a fixed pseudofermion background. Using the Leibniz rule we get $0 = \mathcal{A}(\phi) = \mathcal{A}(\mathcal{M}\mathcal{M}^{-1}\phi) = \mathcal{A}(\mathcal{M})\mathcal{M}^{-1}\phi + \mathcal{M}\mathcal{A}(\mathcal{M}^{-1}\phi)$ so $\mathcal{A}(\mathcal{M}^{-1}\phi) = -\mathcal{M}^{-1}\mathcal{A}(\mathcal{M})\mathcal{M}^{-1}\phi$. To use this for a tower of arbitrary height we generalize this to

$$\begin{aligned} 0 = \mathcal{A}^j(\mathcal{M}\mathcal{M}^{-1}\phi) &= \sum_{k=0}^{j}\binom{j}{k}\mathcal{A}^{j-k}(\mathcal{M})\mathcal{A}^k(\mathcal{M}^{-1}\phi) \\ &= \mathcal{M}\mathcal{A}^j(\mathcal{M}^{-1}\phi) + \sum_{k=0}^{j-1}\binom{j}{k}\mathcal{A}^{j-k}(\mathcal{M})\mathcal{A}^k(\mathcal{M}^{-1}\phi) \end{aligned}$$

for $j > 0$, and thus

$$\mathcal{A}^j(\mathcal{M}^{-1}\phi) = -\mathcal{M}^{-1}\sum_{k=0}^{j-1}\binom{j}{k}\mathcal{A}^{j-k}(\mathcal{M})\mathcal{A}^k(\mathcal{M}^{-1}\phi).$$

This translates into the following recursive definition for the tower on X:

$$\begin{aligned} T(A, X)_0 &= \mathcal{M}^{-1}\phi \\ T(A, X)_j &= -\mathcal{M}^{-1}\sum_{k=0}^{j-1}\binom{j}{k}T(A, \mathcal{M})_{j-k}T(A, X)_k \end{aligned}$$

in terms of the tower $T(A, \mathcal{M})$ which we already know how to compute. Note that we require exactly n inverses to construct such a tower of height of height n.

Yin [29] has suggested an ingenious way of performing a force-gradient update by computing the force twice. We should not be surprised that the force-gradient update $e^{\delta\tau^3\widehat{\{S,\{S,T\}\}}}$ can be computed out of $e^{\delta\tau\hat{S}}$ and $e^{\delta\tau\hat{T}}$ steps:

[9]Here we use the symbol $\cdot$ to emphasize multiplication operations. Elsewhere we use juxtaposition to indicate multiplication.

[10]The symbol $\cdot$ on the left denotes multiplication of towers, whereas on the right it denotes matrix multiplication.

recall that according to the BCH formula the commutator $C(e^A, e^B) = e^{-A}e^{-B}e^A e^B = e^{[A,B]+\cdots}$, hence

$$\begin{aligned} C(e^{\delta\tau\hat{S}}, C(e^{\delta\tau\hat{S}}, e^{\delta\tau\hat{T}})) &= e^{-\delta\tau\hat{S}}e^{-\delta\tau\hat{T}}e^{-\delta\tau\hat{S}}e^{\delta\tau\hat{T}}e^{\delta\tau\hat{S}}e^{-\delta\tau\hat{T}}e^{\delta\tau\hat{S}}e^{\delta\tau\hat{T}} \\ &= C(e^{\delta\tau\hat{S}}, e^{\delta\tau^2[\hat{S},\hat{T}]+\mathcal{O}(\delta\tau^3)}) \\ &= e^{\delta\tau^3[\hat{S},[\hat{S},\hat{T}]]+\mathcal{O}(\delta\tau^4)} \\ &= e^{\delta\tau^3\widehat{\{S,\{S,T\}\}}+\mathcal{O}(\delta\tau^4)}. \end{aligned}$$

It is interesting that this can be reduced to only requiring two inverses in the case where T is quadratic. There does not seem to be a way of using this trick to evaluate Poisson brackets, however.

VII. CONCLUSIONS

We have given a formalism for computing integrators and the corresponding shadow Hamiltonians for lattice gauge theories, and we have presented explicit formulas for the Poisson brackets up to fourth-order and for the force-gradient update step. We have shown how to express these quantities in terms of basic lattice vectors taking their values in the representation of the Lie algebra, as is needed for the usual formulation of lattice gauge theories, and explained how these may be computed using towers. The implementation of towers is straightforward, as it just requires the substitution of the algebra of towers for that of the matrices already used in computing the force term. The stencils for any action are unchanged, and the method is readily applied to pseudofermions, smeared actions, and so forth. The rules for addition, multiplication, and ''inversion'' of towers are given in a recursive form that is easy to implement (although a recursive implementation is not necessary).

ACKNOWLEDGMENTS

We would like to thank Bálint Joó for implementing towers in Chroma. P. J. S. acknowledges support from FCT via Grant No. SFRH/BPD/40998/2007, and Project No. PTDC/FIS/100968/2008, developed under the initiative QREN financed by the UE/FEDER through the Programme COMPETE—''Programa Operacional Factores de Competitividade.''

APPENDIX A: DIFFERENTIAL FORMS

1. Differential forms and wedge products

For convenience we give the definition of a few basic operations on differential forms. In some local basis q: $\mathcal{M} \supseteq \mathcal{U} \to \mathbb{R}^n$ a k-form $\Omega \in \Lambda^k$ has components [11]

$$\begin{aligned} \Omega &= \sum_{1\le I_1<\cdots<I_k\le k} \Omega_{I_1\ldots I_k} \boldsymbol{d}q^{I_1} \wedge \cdots \wedge \boldsymbol{d}q^{I_k} \\ &= \frac{1}{k!}\sum_{i_1,\ldots,i_k=1}^{N} \Omega_{i_1\ldots i_k} \boldsymbol{d}q^{i_1} \wedge \cdots \wedge \boldsymbol{d}q^{i_k} \\ &\equiv \frac{1}{k!}\Omega_{i_1\ldots i_k} \boldsymbol{d}q^{i_1} \wedge \cdots \wedge \boldsymbol{d}q^{i_k} \\ &= \frac{1}{k!}\sum_{\pi\in\mathcal{S}_k} \Omega_{\pi_1\ldots\pi_k} \boldsymbol{d}q^{\pi_1} \wedge \cdots \wedge \boldsymbol{d}q^{\pi_k} \\ &= \langle \Omega_{\pi_1\ldots\pi_k} \boldsymbol{d}q^{\pi_1} \wedge \cdots \wedge \boldsymbol{d}q^{\pi_k} \rangle_{\pi\in\mathcal{S}_k}, \end{aligned}$$

where $\mathcal{S}_k$ is the symmetric group acting on $1, \ldots, k$, and $\langle\cdots\rangle_{\mathcal{S}_k}$ indicates the average over elements of the symmetric group. The *wedge product satisfies*

$\boldsymbol{\alpha} \wedge \boldsymbol{\beta} = (-1)^{kk'}\boldsymbol{\beta} \wedge \boldsymbol{\alpha}$ $\qquad \boldsymbol{\alpha} \in \Lambda^k, \boldsymbol{\beta} \in \Lambda^{k'}$

antisymmetry;

$\boldsymbol{\alpha} \wedge \boldsymbol{\beta} \wedge \boldsymbol{\gamma} = \boldsymbol{\alpha} \wedge (\boldsymbol{\beta} \wedge \boldsymbol{\gamma}) = (\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \wedge \boldsymbol{\gamma}$

associativity.

In terms of the components in local coordinates this means that [12]

$$\begin{aligned} \boldsymbol{\alpha}\wedge\boldsymbol{\beta} &= \langle \alpha_{\pi_1\ldots\pi_k}\beta_{\pi_{k+1}\ldots\pi_{k+k'}} \boldsymbol{d}q^{\pi_1} \wedge \cdots \wedge \boldsymbol{d}q^{\pi_{k+k'}} \rangle_{\pi\in\mathcal{S}_{k+k'}} \\ &= \frac{1}{(k+k')!}\sum_{\pi\in\mathcal{S}_{k+k'}} \alpha_{\pi_1\ldots\pi_k}\beta_{\pi_{k+1}\ldots\pi_{k+k'}} \\ &\quad \times \boldsymbol{d}q^{\pi_1}\wedge\cdots\wedge\boldsymbol{d}q^{\pi_{k+k'}}. \end{aligned}$$

2. Exterior derivatives

The *exterior derivative* $\boldsymbol{d}$: $\Lambda^k \to \Lambda^{k+1}$ is a linear antiderivation, so

$\boldsymbol{d}(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \boldsymbol{d}\boldsymbol{\alpha} + \boldsymbol{d}\boldsymbol{\beta}$ linearity;

$\boldsymbol{d}(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) = (\boldsymbol{d}\boldsymbol{\alpha}) \wedge \boldsymbol{\beta} + (-1)^k\boldsymbol{\alpha} \wedge \boldsymbol{d}\boldsymbol{\beta}$ $\qquad \boldsymbol{\alpha} \in \Lambda^k$

anti-Leibniz;

$\boldsymbol{d}^2\boldsymbol{\alpha} = 0$ $\qquad \boldsymbol{d}F(\boldsymbol{x}) = \boldsymbol{x}F$ $\qquad F \in \Lambda^0$.

The exterior derivative $\boldsymbol{d}F$ for a zero-form F is defined to be $\boldsymbol{d}F(\boldsymbol{x}) \equiv \boldsymbol{x}F$ for any vector field $\boldsymbol{x}$: if we evaluate this in a local coordinate system we find that

$$\begin{aligned} \boldsymbol{d}F(\boldsymbol{x}) = \boldsymbol{x}F &= \left(x^i\frac{\partial}{\partial q^i}\right)F = \left(\frac{\partial F}{\partial q^i}\right)x^i \\ &= \left(\frac{\partial F}{\partial q^i}\boldsymbol{d}q^i\right)\left(x^j\frac{\partial}{\partial q^j}\right) = \left(\frac{\partial F}{\partial q^i}\boldsymbol{d}q^i\right)(\boldsymbol{x}), \end{aligned}$$

so

[11] Our convention is that each independent component occurs once in the sum: another convention is that each such component occurs $k!$ times—once for each permutation of its indices.

[12] For the other convention the numerical coefficient in this formula is $1/(k!k'!)$: *caveat emptor.*

$$\boldsymbol{d}F = \frac{\partial F}{\partial q^i}\boldsymbol{d}q^i.$$

Likewise, in a local coordinate system the exterior derivative of a k-form $\boldsymbol{\Omega} \in \Lambda^k$ is

$$\begin{aligned}\boldsymbol{d}\Omega &= \boldsymbol{d}\left(\frac{1}{k!}\Omega_{i_1\ldots i_k}\boldsymbol{d}q^{i_1}\wedge\cdots\wedge\boldsymbol{d}q^{i_k}\right)\\ &= \frac{1}{k!}\frac{\partial\Omega_{i_1\ldots i_k}}{\partial q^j}\boldsymbol{d}q^j\wedge\boldsymbol{d}q^{i_1}\wedge\cdots\wedge\boldsymbol{d}q^{i_k}.\end{aligned}$$

This follows from the anti-Leibniz rule $\boldsymbol{d}(\alpha\boldsymbol{\beta}) = \boldsymbol{d}\alpha\wedge\boldsymbol{\beta} + \alpha\boldsymbol{d}\boldsymbol{\beta}$ applied to the case where $\alpha = \Omega_{i_1\ldots i_k} \in \Lambda^0$ and $\boldsymbol{\beta} = \boldsymbol{d}q^{i_1}\wedge\cdots\wedge\boldsymbol{d}q^{i_k}$ because the second term vanishes (by induction on k) using the condition $\boldsymbol{d}^2 = 0$ for the basis forms which are exterior derivatives of the coordinates q^i, $\boldsymbol{d}^2q^i = 0$.

In particular, for a one-form $\boldsymbol{\theta} \in \Lambda^1$ we have

$$\boldsymbol{d}\boldsymbol{\theta} = \frac{\partial\theta_i}{\partial q^j}\boldsymbol{d}q^j\wedge\boldsymbol{d}q^i,$$

so applying the two-form $\boldsymbol{d}\boldsymbol{\theta}$ to two arbitrary vector fields $\boldsymbol{x}$ and $\boldsymbol{y}$ gives

$$\begin{aligned}\boldsymbol{d}\boldsymbol{\theta}(\boldsymbol{x},\boldsymbol{y}) &= \frac{\partial\theta_i}{\partial q^j}(x^jy^i - x^iy^j)\\ &= x^j\frac{\partial}{\partial q^j}(\theta_iy^i) - x^j\theta_i\frac{\partial y^i}{\partial q^j} - y^j\frac{\partial}{\partial q^j}(\theta_ix^i) + y^j\theta_i\frac{\partial x^i}{\partial q^j}\\ &= \boldsymbol{x}\boldsymbol{\theta}(\boldsymbol{y}) - \boldsymbol{y}\boldsymbol{\theta}(\boldsymbol{x}) - \theta_i[\boldsymbol{x}(y^i) - \boldsymbol{y}(x^i)]\\ &= \boldsymbol{x}\boldsymbol{\theta}(\boldsymbol{y}) - \boldsymbol{y}\boldsymbol{\theta}(\boldsymbol{x}) - \boldsymbol{\theta}([\boldsymbol{x},\boldsymbol{y}]).\end{aligned} \tag{A1}$$

This provides an elegant coordinate-independent definition of $\boldsymbol{d}\boldsymbol{\theta}$ in terms of the *commutator* of the vector fields

$$\begin{aligned}[\boldsymbol{x},\boldsymbol{y}] \equiv \boldsymbol{x}\boldsymbol{y} - \boldsymbol{y}\boldsymbol{x} &= x^i\frac{\partial}{\partial q^i}y^j\frac{\partial}{\partial q^j} - y^i\frac{\partial}{\partial q^i}x^j\frac{\partial}{\partial q^j}\\ &= \left(x^i\frac{\partial y^j}{\partial q^i} - y^i\frac{\partial x^j}{\partial q^i}\right)\frac{\partial}{\partial q^j} + (x^iy^j - x^jy^i)\frac{\partial}{\partial q^i}\frac{\partial}{\partial q^j},\end{aligned} \tag{A2}$$

which is itself a vector field since the last term involving second derivatives vanishes by symmetry. Note that if $\boldsymbol{\theta}$ is exact, that is $\boldsymbol{\theta} = \boldsymbol{d}F$, then the identity $\boldsymbol{d}^2F(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}\boldsymbol{d}F(\boldsymbol{y}) - \boldsymbol{y}\boldsymbol{d}F(\boldsymbol{x}) - \boldsymbol{d}F([\boldsymbol{x},\boldsymbol{y}]) = \boldsymbol{x}\boldsymbol{y}F - \boldsymbol{y}\boldsymbol{x}F - [\boldsymbol{x},\boldsymbol{y}]F = 0$ holds automatically.

For an arbitrary $(k-1)$-form $\Omega \in \Lambda^{k-1}$ we may derive the corresponding identity,

$$\begin{aligned}&\boldsymbol{d}\Omega(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k)\\ &= \sum_{i=1}^{k}(-1)^{i+1}\boldsymbol{x}_i\Omega(\boldsymbol{x}_1,\ldots,\hat{\boldsymbol{x}}_i,\ldots,\boldsymbol{x}_k)\\ &\quad - \sum_{1\le i<j\le k}(-1)^{i+j+1}\Omega([\boldsymbol{x}_i,\boldsymbol{x}_j],\boldsymbol{x}_1,\ldots,\hat{\boldsymbol{x}}_i,\ldots,\hat{\boldsymbol{x}}_j,\ldots,\boldsymbol{x}_k),\end{aligned}$$

where $\hat{\boldsymbol{x}}$ indicate that the variable $\boldsymbol{x}$ is omitted. We observe that for $k = 3$ the invariant expression for the exterior derivative is

$$\begin{aligned}\boldsymbol{d}\boldsymbol{\omega}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) &= \boldsymbol{x}\boldsymbol{\omega}(\boldsymbol{y},\boldsymbol{z}) - \boldsymbol{y}\boldsymbol{\omega}(\boldsymbol{x},\boldsymbol{z}) + \boldsymbol{z}\boldsymbol{\omega}(\boldsymbol{x},\boldsymbol{y})\\ &\quad - \boldsymbol{\omega}([\boldsymbol{x},\boldsymbol{y}],\boldsymbol{z}) + \boldsymbol{\omega}([\boldsymbol{x},\boldsymbol{z}],\boldsymbol{y}) - \boldsymbol{\omega}([\boldsymbol{y},\boldsymbol{z}],\boldsymbol{x}).\end{aligned} \tag{A3}$$

3. Interior products

The *interior product* i: $T\mathcal{M}\times\Lambda^k \to \Lambda^{k-1}$ is the operation that inserts a vector as the first argument of a k-form to yield a $k-1$-form. It is formally defined by the axioms

$$\begin{aligned}i_{\boldsymbol{x}}(\boldsymbol{\alpha}+\boldsymbol{\beta}) &= i_{\boldsymbol{x}}\boldsymbol{\alpha} + i_{\boldsymbol{x}}\boldsymbol{\beta} \qquad \boldsymbol{\alpha},\boldsymbol{\beta}\in\Lambda^k \quad \text{linearity};\\ i_{\boldsymbol{x}}(\boldsymbol{\alpha}\wedge\boldsymbol{\beta}) &= i_{\boldsymbol{x}}(\boldsymbol{\alpha})\wedge\boldsymbol{\beta} + (-1)^k\boldsymbol{\alpha}\wedge i_{\boldsymbol{x}}\boldsymbol{\beta}\\ &\qquad \boldsymbol{\alpha}\in\Lambda^k,\boldsymbol{\beta}\in\Lambda^{k'} \quad \text{anti-Leibniz};\\ i_{\boldsymbol{x}}F &= 0 \qquad F\in\Lambda^0\\ i_{\boldsymbol{x}}\Omega(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{k-1}) &= \Omega(\boldsymbol{x},\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{k-1}) \qquad \Omega\in\Lambda^k;\\ i_{\boldsymbol{x}}^2 &= 0\end{aligned}$$

so we see that it too is a linear antiderivation.

4. Induced Maps

If σ: $\mathcal{M}\to\mathcal{M}'$ is a diffeomorphism, then there is a natural induced map σ_*: $\Lambda^0(\mathcal{M}')\to\Lambda^0(\mathcal{M})$ defined by σ_*f: $p\mapsto f(\sigma p)$ for all $f\in\Lambda^0(\mathcal{M}')$ and $p\in\mathcal{M}$. This map may also be written as $\sigma_*f = f\circ\sigma$ and is called a *pull-back*. Another way of saying this is that the following diagram commutes:

$$\begin{array}{ccc} M & \xrightarrow{\sigma} & \mathcal{M}'\\ {\scriptstyle\sigma_*f}\searrow & & \swarrow{\scriptstyle f}\\ & \mathbb{R} & \end{array}$$

If σ^{-1} exists then there is a corresponding pull-back map $(\sigma^{-1})_*$, and it satisfies the relation $(\sigma^{-1})_*\sigma_*f = (\sigma^{-1})_*(f\circ\sigma) = f\circ\sigma\circ\sigma^{-1} = f$, and thus we see that $(\sigma^{-1})_* = (\sigma_*)^{-1}$, and we may denote both of these unambiguously as σ_*^{-1}.

If $\boldsymbol{x}\in T\mathcal{M}$ is a vector field on $\mathcal{M}$ then there may be a *push-through* map σ^*: $T\mathcal{M}\to T\mathcal{M}'$ defined by $\sigma^*\boldsymbol{x} = \sigma_*^{-1}\circ\boldsymbol{x}\circ\sigma_*$ if this exists. For any $f\in\Lambda^0(\mathcal{M}')$ and $p\in\mathcal{M}$ this means that $\sigma^*\boldsymbol{x}(f)|_{\sigma p} = \boldsymbol{x}(\sigma_*f)|_p$. The corresponding commutative diagram is

$$\begin{array}{ccc} \Lambda^0(\mathcal{M}) & \xleftarrow{\sigma_*} & \Lambda^0(\mathcal{M}')\\ \boldsymbol{x}\downarrow & & \downarrow\sigma^*\boldsymbol{x}\\ \Lambda^0(\mathcal{M}) & \xrightarrow{\sigma_*^{-1}} & \Lambda^0(\mathcal{M}') \end{array}.$$

The existence of the diffeomorphism σ^{-1}: $\mathcal{M}'\to\mathcal{M}$ is a sufficient but not necessary condition for σ_*^{-1} and hence σ^* to be well defined.

We may define further induced maps [13] such as the pull-back of one-form fields $\sigma_*\colon \Lambda^1(\mathcal{M}') \to \Lambda^1(\mathcal{M})$ as $\sigma_*\boldsymbol{\theta} = \sigma_* \circ \boldsymbol{\theta} \circ \sigma^*$,

$$\begin{array}{ccc} T\mathcal{M} & \overset{\sigma^*}{\rightarrow} & T\mathcal{M}' \\ \sigma_*\boldsymbol{\theta} \downarrow & & \downarrow \boldsymbol{\theta} \\ \Lambda^0(\mathcal{M}) & \overset{\sigma_*}{\leftarrow} & \Lambda^0(\mathcal{M}'), \end{array}$$

and so forth.

In the special case where $\sigma\colon \mathcal{M} \to \mathcal{M}$ is an auto-diffeomorphism then the push-through maps always exist.

5. Lie derivatives

Suppose now that we have a smooth one-parameter family of diffeomorphisms $\sigma\colon \mathbb{R} \times \mathcal{M} \to \mathcal{M}$, which we will also write as $\sigma_t\colon \mathcal{M} \to \mathcal{M}$. Using this map we can define a derivative with respect to the parameter t, which is called a *Lie derivative*. For any zero-form F we define

$$\mathcal{L}_{\boldsymbol{v}} F \equiv \frac{d(\sigma_{t*}F)}{dt}\bigg|_{t=0} = \frac{d(F \circ \sigma_t)}{dt}\bigg|_{t=0} = \boldsymbol{v}F, \qquad \text{(A4)}$$

where $\boldsymbol{v}$ is the linear differential operator—the vector field—that is tangential to the curves $\sigma(t, p)$ passing through $\sigma(0, p) = p \in \mathcal{M}$ at $t = 0$.

The Lie derivative of a vector field $\boldsymbol{y} \in T\mathcal{M}$ can be deduced from the requirement that $\mathcal{L}_{\boldsymbol{x}}$ be a derivation

$$\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{A} \otimes \boldsymbol{B}) = (\mathcal{L}_{\boldsymbol{x}}\boldsymbol{A}) \otimes \boldsymbol{B} + \boldsymbol{A} \otimes \mathcal{L}_{\boldsymbol{x}}\boldsymbol{B}$$

for any tensors $\boldsymbol{A}$ and $\boldsymbol{B}$, and that it commutes with contractions

$$\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{y}F) = (\mathcal{L}_{\boldsymbol{x}}\boldsymbol{y})F + \boldsymbol{y}(\mathcal{L}_{\boldsymbol{x}}F),$$
$$\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{\theta}(\boldsymbol{y})) = (\mathcal{L}_{\boldsymbol{x}}\boldsymbol{\theta})(\boldsymbol{y}) + \boldsymbol{\theta}(\mathcal{L}_{\boldsymbol{x}}\boldsymbol{y}),$$

and so forth. Applying these rules to the zero-form $\boldsymbol{y}F$ obtained by applying the vector field $\boldsymbol{y} \in T\mathcal{M}$ to $F \in \Lambda^0(\mathcal{M})$ we have $\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{y}F) = \boldsymbol{x}\boldsymbol{y}F$ and also $\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{y}F) = (\mathcal{L}_{\boldsymbol{x}}\boldsymbol{y})F + \boldsymbol{y}(\mathcal{L}_{\boldsymbol{x}}F)$, hence

$$(\mathcal{L}_{\boldsymbol{x}}\boldsymbol{y})F = \boldsymbol{x}\boldsymbol{y}F - \boldsymbol{y}\boldsymbol{x}F = [\boldsymbol{x}, \boldsymbol{y}]F$$

and, as this holds for all F,

$$\mathcal{L}_{\boldsymbol{x}}\boldsymbol{y} = [\boldsymbol{x}, \boldsymbol{y}]. \qquad \text{(A5)}$$

We may apply a similar argument to evaluate the Lie derivative of a one-form $\boldsymbol{\theta} \in \Lambda^1(\mathcal{M})$. On the one hand $\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{\theta}(\boldsymbol{y})) = \boldsymbol{x}\boldsymbol{\theta}(\boldsymbol{y})$, while on the other $\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{\theta}(\boldsymbol{y})) = (\mathcal{L}_{\boldsymbol{x}}\boldsymbol{\theta})(\boldsymbol{y}) + \boldsymbol{\theta}(\mathcal{L}_{\boldsymbol{x}}\boldsymbol{y})$, so using (A1)

$$\begin{aligned}(\mathcal{L}_{\boldsymbol{x}}\boldsymbol{\theta})(\boldsymbol{y}) &= \boldsymbol{x}\boldsymbol{\theta}(\boldsymbol{y}) - \boldsymbol{\theta}([\boldsymbol{x}, \boldsymbol{y}]) = \boldsymbol{d}\boldsymbol{\theta}(\boldsymbol{x}, \boldsymbol{y}) + \boldsymbol{y}\boldsymbol{\theta}(\boldsymbol{x}) \\ &= (i_{\boldsymbol{x}}\boldsymbol{d}\boldsymbol{\theta})(\boldsymbol{y}) + \boldsymbol{d}(\boldsymbol{\theta}(\boldsymbol{x}))(\boldsymbol{y}) \\ &= (i_{\boldsymbol{x}}\boldsymbol{d}\boldsymbol{\theta})(\boldsymbol{y}) + (\boldsymbol{d}i_{\boldsymbol{x}}\boldsymbol{\theta})(\boldsymbol{y}) = (i_{\boldsymbol{x}}\boldsymbol{d} + \boldsymbol{d}i_{\boldsymbol{x}})\boldsymbol{\theta}(\boldsymbol{y}),\end{aligned}$$

hence

$$\mathcal{L}_{\boldsymbol{x}}\boldsymbol{\theta} = (i_{\boldsymbol{x}}\boldsymbol{d} + \boldsymbol{d}i_{\boldsymbol{x}})\boldsymbol{\theta}.$$

This suggests that the Lie derivative of any k-form may be expressed as

$$\mathcal{L}_{\boldsymbol{x}} = i_{\boldsymbol{x}}\boldsymbol{d} + \boldsymbol{d}i_{\boldsymbol{x}}, \qquad \text{(A6)}$$

and this is indeed the case as the operator $i_{\boldsymbol{x}}\boldsymbol{d} + \boldsymbol{d}i_{\boldsymbol{x}}$ is a derivation

$$\begin{aligned}(i_{\boldsymbol{x}}\boldsymbol{d} + \boldsymbol{d}i_{\boldsymbol{x}})(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) &= i_{\boldsymbol{x}}[(\boldsymbol{d}\boldsymbol{\alpha}) \wedge \boldsymbol{\beta} + (-1)^k\boldsymbol{\alpha} \wedge \boldsymbol{d}\boldsymbol{\beta}] + \boldsymbol{d}[(i_{\boldsymbol{x}}\boldsymbol{\alpha}) \wedge \boldsymbol{\beta} + (-1)^k\boldsymbol{\alpha} \wedge i_{\boldsymbol{x}}\boldsymbol{\beta}] \\ &= (i_{\boldsymbol{x}}\boldsymbol{d}\boldsymbol{\alpha}) \wedge \boldsymbol{\beta} + (-1)^{k+1}(\boldsymbol{d}\boldsymbol{\alpha}) \wedge i_{\boldsymbol{x}}\boldsymbol{\beta} + (-1)^k(i_{\boldsymbol{x}}\boldsymbol{\alpha}) \wedge \boldsymbol{d}\boldsymbol{\beta} + (-1)^{2k}\boldsymbol{\alpha} \wedge i_{\boldsymbol{x}}\boldsymbol{d}\boldsymbol{\beta} + (\boldsymbol{d}i_{\boldsymbol{x}}\boldsymbol{\alpha}) \wedge \boldsymbol{\beta} \\ &\quad + (-1)^{k-1}(i_{\boldsymbol{x}}\boldsymbol{\alpha}) \wedge \boldsymbol{d}\boldsymbol{\beta} + (-1)^k(\boldsymbol{d}\boldsymbol{\alpha}) \wedge i_{\boldsymbol{x}}\boldsymbol{\beta} + (-1)^{2k}\boldsymbol{\alpha} \wedge \boldsymbol{d}i_{\boldsymbol{x}}\boldsymbol{\beta} \\ &= [(i_{\boldsymbol{x}}\boldsymbol{d} + \boldsymbol{d}i_{\boldsymbol{x}})\boldsymbol{\alpha}] \wedge \boldsymbol{\beta} + \boldsymbol{\alpha} \wedge (i_{\boldsymbol{x}}\boldsymbol{d} + \boldsymbol{d}i_{\boldsymbol{x}})\boldsymbol{\beta}\end{aligned}$$

for all $\boldsymbol{\alpha} \in \Lambda^k$ and $\boldsymbol{\beta} \in \Lambda^{k'}$, and for zero- and one-forms F and $\boldsymbol{\theta}$

$$\mathcal{L}_{\boldsymbol{x}}F = \boldsymbol{x}F = \boldsymbol{d}F(\boldsymbol{x}) = i_{\boldsymbol{x}}\boldsymbol{d}F + \boldsymbol{d}i_{\boldsymbol{x}}F,$$
$$\mathcal{L}_{\boldsymbol{x}}\boldsymbol{\theta} = (i_{\boldsymbol{x}}\boldsymbol{d} + \boldsymbol{d}i_{\boldsymbol{x}})\boldsymbol{\theta}.$$

The second term in the first equation is zero because $i_{\boldsymbol{x}}F = 0$ by definition.

[13] One must be careful with the notation introduced here, as there are a whole family of mappings that we have given the same name, $\sigma_*\colon \Lambda^k(\mathcal{M}') \to \Lambda^k(\mathcal{M}) \quad \forall\ k$, and the equation $\sigma_*\boldsymbol{\theta} = \sigma_* \circ \boldsymbol{\theta} \circ \sigma^*$ involves two of them. If we were to call these induced mappings on forms $\sigma^k_*\colon \Lambda^k(\mathcal{M}') \to \Lambda^k(\mathcal{M})$ then the equation is less ambiguous, $\sigma^1_*\boldsymbol{\theta} = \sigma^0_* \circ \boldsymbol{\theta} \circ \sigma^*$.

APPENDIX B: LIE GROUPS

1. Left-invariant forms

A *Lie group* is a manifold that has a group structure defined by C^∞ multiplication $(g, h) \mapsto gh$ and inverse $g \mapsto g^{-1}$ operations that satisfy the group axioms

$$g(g'g'') = (gg')g'' \equiv gg'g'' \quad \forall\ g, g', g'' \in \mathcal{G}$$

associative

$$g^{-1}g = gg^{-1} = \mathbb{I} \quad \forall\ g \in \mathcal{G} \quad \text{inverse}$$

with $\mathbb{I}$ being the identity element of the group. If we consider a point $g \in \mathcal{G}$ as being "fixed" then left multiplication by g is an autodiffeomorphism of $\mathcal{G}$,

L_g: $g' \mapsto gg'$, with $L_{gh} = L_g \circ L_h$ by associativity, $L_g \circ L_h g' = g(hg') = (gh)g' = L_{gh}g'$ for all $g' \in \mathcal{G}$. Clearly $L_{g^{-1}} = (L_g)^{-1}$ too.

As for any such diffeomorphisms we can define the corresponding pull-back maps on forms and vectors, $L_{g*}F \equiv F \circ L_g$, $L_g^*\boldsymbol{v} \equiv L_{g^{-1}*} \circ \boldsymbol{v} \circ L_{g*}$, and $L_{g*}\boldsymbol{\theta} \equiv L_{g*} \circ \boldsymbol{\theta} \circ L_g^*$. We may use these maps to define *left-invariant* vector fields and forms; for example, a left-invariant one-form satisfies the condition $\boldsymbol{\theta} = L_{g*}\boldsymbol{\theta}$.

2. Lie algebra

The only left-invariant zero-forms are constants, as if $F = L_{g*}F$ $(\forall\ g \in \mathcal{G})$ then $F(g) = F(L_g\mathbb{I}) = L_{g*}F(\mathbb{I}) = F(\mathbb{I})$.

If $\boldsymbol{u} = L_g^*\boldsymbol{u}$ and $\boldsymbol{v} = L_g^*\boldsymbol{v}$ are left-invariant vector fields in the tangent bundle $T\mathcal{G}$ then their commutator is also a vector field, and furthermore it is also left-invariant since[14] $[\boldsymbol{u}, \boldsymbol{v}] = [L_g^*\boldsymbol{u}, L_g^*\boldsymbol{v}] = [L_{g^{-1}*} \circ \boldsymbol{u} \circ L_{g*}, L_{g^{-1}*} \circ \boldsymbol{v} \circ L_{g*}] = L_{g^{-1}*} \circ [\boldsymbol{u}, \boldsymbol{v}] \circ L_{g*} = L_g^*[\boldsymbol{u}, \boldsymbol{v}]$. If a left-invariant vector field $\boldsymbol{v}$ vanishes at the identity, $\boldsymbol{v}(F)|_{\mathbb{I}} = 0 (\forall F \in \Lambda^0\mathcal{G})$, then it must be identically zero everywhere, as $\boldsymbol{v}(F)|_g = [\boldsymbol{v}(F) \circ L_g]_{\mathbb{I}} = [L_{g*}\boldsymbol{v}(F)]_{\mathbb{I}} = [L_{g*} \circ L_g^*\boldsymbol{v}(F)]_{\mathbb{I}} = [L_{g*} \circ L_{g*}^{-1} \circ \boldsymbol{v}(L_{g*}F)]_{\mathbb{I}} = [\boldsymbol{v}(F \circ L_g)]_{\mathbb{I}} = 0$.

Consider a set of left-invariant vector fields $\{\boldsymbol{e}_i\}$ in $T\mathcal{G}$ called *generators* whose values at the origin are linearly independent; any linear combination of the generators with left-invariant (constant) coefficients is also left-invariant. Conversely any left-invariant vector field $\boldsymbol{u}$ must be a linear combination of this type, since its value at the origin is $\boldsymbol{u}|_{\mathbb{I}} = \sum_i u^i \boldsymbol{e}_i|_{\mathbb{I}}$ with $u^i \in \mathbb{R}$, and hence $\boldsymbol{u} - \sum_i u^i \boldsymbol{e}_i = 0$ everywhere. Left-invariant vector fields therefore form a linear space with constant coefficients. In particular, the commutator of any left-invariant vector fields must be a linear combination of the generators, $[\boldsymbol{e}_i, \boldsymbol{e}_j] = c_{ij}^k \boldsymbol{e}_k$ where the $c_{ij}^k \in \mathbb{R}$ are called the *structure constants*. This makes the linear space of left-invariant vector fields into a *Lie algebra*.

Any left-invariant vector field $\boldsymbol{v}$ has an integral curve[15] $c: \mathbb{R} \to \mathcal{G}$ satisfying $c(0) = \mathbb{I}$. Along this curve we have an Abelian subgroup of $\mathcal{G}$ satisfying $c(s+t) = c(s)c(t)$, so it is naturally to call c an *exponential map*, and write it as $c(t) = \exp(\boldsymbol{v}t)$. If we view this as a function of $\boldsymbol{v}$ then this defines a *local flow* of $\boldsymbol{v}$, and is a map from the Lie algebra into the Lie group, exp: $T\mathcal{G} \to \mathcal{G}$.

The *commutator* of two elements $g, h \in \mathcal{G}$ is defined to be $C(g.h) \equiv g^{-1}h^{-1}gh$; in a neighborhood of the identity where $g = \exp(\boldsymbol{u}t)$ and $h = \exp(\boldsymbol{v}t)$ we have

$$
\begin{aligned}
C(g,h) &= \exp(-\boldsymbol{u}t)\exp(-\boldsymbol{v}t)\exp(\boldsymbol{u}t)\exp(\boldsymbol{v}t) \\
&= \left(\mathbb{I} - \boldsymbol{u}t + \frac{1}{2}(\boldsymbol{u}t)^2\right)\left(\mathbb{I} - \boldsymbol{v}t + \frac{1}{2}(\boldsymbol{v}t)^2\right) \\
&\quad \times \left(\mathbb{I} + \boldsymbol{u}t + \frac{1}{2}(\boldsymbol{u}t)^2\right)\left(\mathbb{I} + \boldsymbol{v}t + \frac{1}{2}(\boldsymbol{v}t)^2\right) + \mathcal{O}(t^3) \\
&= \mathbb{I} + [\boldsymbol{u}, \boldsymbol{v}]t^2 + \mathcal{O}(t^3) = \exp([\boldsymbol{u}, \boldsymbol{v}]t^2) + \mathcal{O}(t^3).
\end{aligned}
$$

3. Maurer-Cartan equations

The commutation relations may be succinctly expressed in terms of the cotangent space $T^*\mathcal{G}$. We introduce a set of left-invariant 1-forms $\boldsymbol{\theta}^i$ (called a frame or *repère mobile*) dual to the generators $\boldsymbol{\theta}^i(\boldsymbol{e}_j) = \delta_j^i$. From (A1) we have

$$
\begin{aligned}
\boldsymbol{d\theta}^i(\boldsymbol{e}_j, \boldsymbol{e}_k) &= \boldsymbol{e}_j\boldsymbol{\theta}^i(\boldsymbol{e}_k) - \boldsymbol{e}_k\boldsymbol{\theta}^i(\boldsymbol{e}_j) - \boldsymbol{\theta}^i([\boldsymbol{e}_j, \boldsymbol{e}_k]) \\
&= \boldsymbol{e}_j\delta_k^i - \boldsymbol{e}_k\delta_j^i - \boldsymbol{\theta}^i(c_{jk}^\ell \boldsymbol{e}_\ell) \\
&= -c_{jk}^\ell \delta_\ell^i = -c_{jk}^i,
\end{aligned}
$$

so expanding the two-form $\boldsymbol{d\theta}^i = \alpha_{mn}^i \boldsymbol{\theta}^m \wedge \boldsymbol{\theta}^n$ in terms of the basis two-forms $\boldsymbol{\theta}^m \wedge \boldsymbol{\theta}^n$ we have

$$
\begin{aligned}
\boldsymbol{d\theta}(\boldsymbol{e}_j, \boldsymbol{e}_k) &= \alpha_{mn}^i \boldsymbol{\theta}^m \wedge \boldsymbol{\theta}^n(\boldsymbol{e}_j, \boldsymbol{e}_k) \\
&= \alpha_{mn}^i \{\boldsymbol{\theta}^m(\boldsymbol{e}_j)\boldsymbol{\theta}^n(\boldsymbol{e}_k) - \boldsymbol{\theta}^m(\boldsymbol{e}_k)\boldsymbol{\theta}^n(\boldsymbol{e}_j)\} \\
&= \alpha_{mn}^i \{\delta_j^m\delta_k^n - \delta_k^m\delta_j^n\} = \alpha_{jk}^i - \alpha_{kj}^i,
\end{aligned}
$$

thus the left-invariant forms $\boldsymbol{\theta}^i$ satisfy the Maurer-Cartan equations $d\boldsymbol{\theta}^i = -\frac{1}{2}c_{jk}^i \boldsymbol{\theta}^j \wedge \boldsymbol{\theta}^k$ everywhere.

4. Adjoint representation

For any Lie algebra the *adjoint representation* is defined by $\mathrm{ad}(\boldsymbol{x})\boldsymbol{y} = [\boldsymbol{x}, \boldsymbol{y}]$. This is a representation of the Lie algebra because for any $\boldsymbol{z}$

$$
\begin{aligned}
[\mathrm{ad}(\boldsymbol{x}), \mathrm{ad}(\boldsymbol{y})]\boldsymbol{z} &= \mathrm{ad}(\boldsymbol{x})\mathrm{ad}(\boldsymbol{y})\boldsymbol{z} - \mathrm{ad}(\boldsymbol{y})\mathrm{ad}(\boldsymbol{x})\boldsymbol{z} \\
&= \mathrm{ad}(\boldsymbol{x})[\boldsymbol{y}, \boldsymbol{z}] - \mathrm{ad}(\boldsymbol{y})[\boldsymbol{x}, \boldsymbol{z}] \\
&= [\boldsymbol{x}, [\boldsymbol{y}, \boldsymbol{z}]] - [\boldsymbol{y}, [\boldsymbol{x}, \boldsymbol{z}]] \\
&= [[\boldsymbol{x}, \boldsymbol{y}], \boldsymbol{z}] = \mathrm{ad}([\boldsymbol{x}, \boldsymbol{y}])\boldsymbol{z},
\end{aligned}
$$

where we used the Jacobi identity in the penultimate step, and thus $[\mathrm{ad}(\boldsymbol{x}), \mathrm{ad}(\boldsymbol{y})] = \mathrm{ad}([\boldsymbol{x}, \boldsymbol{y}])$. In terms of basis vectors we have $\mathrm{ad}(\boldsymbol{e}_i)\boldsymbol{e}_j = [\boldsymbol{e}_i, \boldsymbol{e}_j] = c_{ij}^k \boldsymbol{e}_k$, giving the explicit matrices $\mathrm{ad}(\boldsymbol{e}_i)_j^k = c_{ij}^k$.

5. Cartan-Killing metric

We may use the adjoint representation to define the *Cartan-Killing metric* on the Lie algebra as a trace, $\langle \boldsymbol{x}, \boldsymbol{y} \rangle \equiv \mathrm{tr}[\mathrm{ad}(\boldsymbol{x})\mathrm{ad}(\boldsymbol{y})]/C_A$ where C_A is a constant; in terms of the basis vectors $g_{ij} \equiv \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \mathrm{tr}[\mathrm{ad}(\boldsymbol{e}_i)\mathrm{ad}(\boldsymbol{e}_j)]/C_A = c_{il}^k c_{jk}^l / C_A$. For a *semisimple* Lie algebra the Cartan-Killing metric is nonsingular and has an inverse satisfying $g^{ij}g_{jk} = \delta_k^i$. For a *simple* Lie algebra

[14] Note that $L_{g^{-1}*} = (L_g)_*^{-1}$.

[15] Strictly speaking this is only true locally: to be precise we should write $c: \mathcal{I} \to \mathcal{M}$ where $\mathcal{I} \subseteq \mathbb{R}$ is a neighborhood of zero.

the adjoint representation is irreducible, so by Schur's lemma the invariant Cartan-Killing metric is a multiple of the unit matrix; we shall choose the constant C_A such that this multiple is unity. For su(N) where the generators in the defining N dimensional fundamental representation T_i satisfy the commutation relations $[T_i, T_j] = c_{ij}^k T_k$ and are normalized such that $\mathrm{tr} T_i T_j = a\delta_{ij}$ the Cartan-Killing metric is explicitly $g_{ij} = \delta_{ij}$ with $C_A = 2aN$.

For semisimple Lie algebras we can use the Cartan-Killing metric and its inverse to lower and raise indices at will, for example, we shall define $p^i \equiv g^{ij} p_i$, and correspondingly we have an invariant quadratic form for one-forms, $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = g^{ij}\alpha_i\beta_j$ where $\boldsymbol{\alpha} = \alpha_i\theta^i$ and $\boldsymbol{\beta} = \beta_i\theta^i$. We also note that the quantity $c_{ijk} = g_{i\ell}c_{jk}^{\ell} = -c_{ikj}$ is totally antisymmetric, because $\langle [\boldsymbol{e}_i, \boldsymbol{e}_j], \boldsymbol{e}_k \rangle = c_{ij}^{\ell}\langle \boldsymbol{e}_\ell, \boldsymbol{e}_k \rangle = c_{ij}^{\ell} g_{\ell k} = c_{kij}$, and

$$
\begin{aligned}
C_A\langle [\boldsymbol{X}, \boldsymbol{Y}], \boldsymbol{Z} \rangle &= \mathrm{tr}(\mathrm{ad}([\boldsymbol{X}, \boldsymbol{Y}])\mathrm{ad}(\boldsymbol{Z})) \\
&= \mathrm{tr}([\mathrm{ad}(\boldsymbol{X}), \mathrm{ad}(\boldsymbol{Y})]\mathrm{ad}(\boldsymbol{Z})) \\
&= \mathrm{tr}(\mathrm{ad}(\boldsymbol{X})\mathrm{ad}(\boldsymbol{Y})\mathrm{ad}(\boldsymbol{Z}) - \mathrm{ad}(\boldsymbol{Y})\mathrm{ad}(\boldsymbol{X})\mathrm{ad}(\boldsymbol{Z})) \\
&= \mathrm{tr}(\mathrm{ad}(\boldsymbol{Z})\mathrm{ad}(\boldsymbol{X})\mathrm{ad}(\boldsymbol{Y}) - \mathrm{ad}(\boldsymbol{X})\mathrm{ad}(\boldsymbol{Z})\mathrm{ad}(\boldsymbol{Y})) \\
&= \mathrm{tr}([\mathrm{ad}(\boldsymbol{Z}), \mathrm{ad}(\boldsymbol{X})]\mathrm{ad}(\boldsymbol{Y})) \\
&= \mathrm{tr}(\mathrm{ad}([\boldsymbol{Z}, \boldsymbol{X}])\mathrm{ad}(\boldsymbol{Y})) \\
&= C_A\langle [\boldsymbol{Z}, \boldsymbol{X}], \boldsymbol{Y} \rangle,
\end{aligned}
$$

hence $c_{ijk} = c_{jki} = c_{kij}$.

[1] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).
[2] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration* (Springer, New York, 2006), 2nd ed.
[3] I. P. Omelyan, I. M. Mryglod, and R. Folk, Phys. Rev. Lett. **86**, 898 (2001).
[4] T. Takaishi, Comput. Phys. Commun. **133**, 6 (2000).
[5] T. Takaishi and P. de Forcrand, Phys. Rev. E **73**, 036706 (2006).
[6] I. T. Drummond, S. Duane, and R. R. Horgan, Nucl. Phys. **B220**, 119 (1983).
[7] A. D. Kennedy and P. Rossi, Nucl. Phys. **B327**, 782 (1989).
[8] R. L. Bishop and S. I. Goldberg, *Tensor Analysis on Manifolds* (Dover, New York, 1980).
[9] *Analysis, Manifolds and Physics*, edited by Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick (North-Holland, Amsterdam, 1977).
[10] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces* (Academic Press, New York, 1978).
[11] N. J. Hicks, *Notes on Differential Geometry* (Van Nostrand Reinhold, New York, 1971).
[12] M. Spivak, *Differential Geometry* (Publish or Perish, Inc., Berkeley, CA, 1970).
[13] I. P. Omelyan, I. M. Mryglod, and R. Folk, Phys. Rev. E **66**, 026701 (2002).
[14] I. P. Omelyan, I. M. Mryglod, and R. Folk, Comput. Phys. Commun. **151**, 273 (2003).
[15] M. A. Clark, A. D. Kennedy, and P. J. Silva, Proc. Sci. LAT2008 (2008) 041.
[16] M. A. Clark, B. Joó, A. D. Kennedy, and P. J. Silva, Proc. Sci. LATTICE2010 (2010) 323.
[17] M. A. Clark, B. Joó, A. D. Kennedy, and P. J. Silva, Phys. Rev. D **84**, 071502 (2011).
[18] A. D. Kennedy, M. A. Clark, and P. J. Silva, Proc. Sci. LAT2009 (2009) 021.
[19] R. Morrin, A. O'Cais, M. Peardon, S. M. Ryan, and J.-I. Skullerud, Phys. Rev. D **74**, 014505 (2006).
[20] H.-W. Lin *et al.* (Hadron Spectrum Collaboration), Phys. Rev. D **79**, 034502 (2009).
[21] M. Hasenbusch, Phys. Lett. B **519**, 177 (2001).
[22] M. Hasenbusch and K. Jansen, Nucl. Phys. **B659**, 299 (2003).
[23] M. Lüscher, Comput. Phys. Commun. **165**, 199 (2005).
[24] M. A. Clark and A. D. Kennedy, Phys. Rev. Lett. **98**, 051601 (2007).
[25] M. A. Clark and A. D. Kennedy, Phys. Rev. D **75**, 011502 (2007).
[26] R. Berndt, *An Introduction to Symplectic Geometry* (American Mathematical Society, Providence, 2001), Vol. 21.
[27] M. A. Clark and A. D. Kennedy, Phys. Rev. D **76**, 074508 (2007).
[28] S. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, Phys. Rev. D **35**, 2531 (1987).
[29] H. Yin and R. D. Mawhinney, Proc. Sci. LATTICE2011 (2011) 051.