

Investigating the Hybrid Monte Carlo with coherent conjugate momenta

Robert D. Mawhinney^{a,*}

^a*Department of Physics, Columbia University,
New York, New York 10027 U.S.A.*

E-mail: rdm10@columbia.edu

In this talk, we discuss tests of the Hybrid Monte Carlo algorithm using four dimensional pure SU(3) gauge theory when the conjugate momenta are not chosen as random Gaussian variables of uniform variance for each lattice site, but instead are represented as different normal modes across the lattice volume, with variable variance. Generically, this involves simulating in a fixed gauge. One goal of this work is to investigate whether appropriate conjugate momentum modes can, for example, speed up a particular part of the evolution of the lattices, such as topological charge.

*The 42nd International Symposium on Lattice Field Theory (LATTICE2025)
2-8 November 2025
Tata Institute of Fundamental Research, Mumbai, India*

*Speaker

1. Introduction

For simulations of QCD with quarks, the dominant algorithm is the Hybrid Monte Carlo (HMC) [1] or the Rational Hybrid Monte Carlo (RHMC) [2], if roots of squared Hermitian Dirac operator $D^\dagger D$ are needed. These algorithms move the gauge and pseudofermion degrees of freedom through the phase space of the Euclidean Feynman Path Integral by using the original Euclidean gauge action as a potential, defining a set of conjugate momenta for the coordinates in this potential, randomly drawing initial pseudofermion variables and evolving this entire system as a Hamiltonian with the introduction of a new time, the molecular dynamics time. By controlling the finite time step used in the numerical integration of Hamilton's equations, the change in the Hamiltonian along a molecular dynamics trajectory can be made small enough for the new state of the system at the end of a trajectory to be accepted a reasonable amount of the time when a Metropolis accept/reject step is applied.

While these algorithms are exact, meaning that they sample the phase space under consideration without error, up to the numerical accuracy of the Metropolis accept/reject step, the efficiency with which they move through phase space is of major practical importance. For the gauge fields in QCD, the high energy modes are approaching free modes due to asymptotic freedom, while the low energy modes, which have important non-linear interactions, control the low energy, non-perturbative physics. Ideally, one would like to move more quickly through the phase space of low-energy modes, but the standard (R)HMC does not allow this balance to be adjusted. From here on, we focus on the evolution of the gauge fields in the absence of fermions.

A number of methods have been proposed to change the balance between low and high mode evolution within the framework of the HMC, and this is an active area of research. Two methods to change this balance are the Riemannian Manifold Hybrid Monte Carlo (RMHMC) [3, 4] and the Field Transformation HMC [5]. In this note, we take a different approach by investigating different choices for the conjugate momenta at the beginning of each trajectory. In the standard HMC, the conjugate momenta are drawn from a Gaussian distribution, which means for a pure gauge theory, conjugate momentum are randomly Gaussianly distributed for each link at each spatial point in Euclidean space, with one momentum for every element of the Lie algebra of $SU(N)$ on each link, *i.e.* 8 for $SU(3)$. We will refer to this as local conjugate momenta.

In this work, we investigate the effects of two changes in the conjugate momentum draw: 1) at the start of each trajectory, we make the width of the Gaussian for each conjugate momentum random, with the constraint that the average width over all conjugate momenta is $1/2$, in the normalization we use, and 2) we consider orthogonal modes for the conjugate momenta which are coherent over the spatial and color degrees of freedom. This coherence, along with the probability that for some trajectories a particular coherent mode will have a larger amplitude, makes it possible that there could be a net improvement in phase-space sampling.

Because of the local gauge invariance of the system, the evolution will have to be done in a fixed gauge for the coherence of the conjugate momenta to have any effect. We add to this the general understanding that low modes in QCD are coherent over ≈ 0.5 fm scales and we are led to be most interested in imagining using this approach for small subvolumes (4^4 to 6^4 at currently accessible lattice spacings). Thus we would not need to gauge fix over the entire lattice volume, but could instead independently gauge fix in, say 4^4 , subvolumes and try different types of coherent

modes within that subvolume.

The organization of this paper is as follows. First we discuss the coherent conjugate momenta needed to create an approximate instanton on the lattice, when no other degrees of freedom are included. This serves as an example of one type of coherent conjugate momenta that might be considered, but the general approach does not restrict the orthogonal modes used to describe the coherent conjugate momenta. The second section presents some simple results concerning the role of the trajectory length on the autocorrelation time of the HMC. Since we will present results for a modified version of the HMC, we wanted to look carefully at some baseline results for the standard HMC. Next we discuss the effects of using local conjugate momenta, but with variable width for their random number draw. Then we will investigate the HMC on a small, 4^4 lattice, where the evolution is done in axial gauge and we use coherent conjugate momenta defined by Fourier transforming to real space conjugate momenta that are local in momentum space.

2. Free Field Example: Coherent Conjugate Momenta Changing Topology

We first consider whether we can produce a localized, instanton-like configuration on the lattice through a molecular dynamics (MD) trajectory that starts with all links equal to one and a particular configuration of conjugate momenta. If we can find such a configuration of conjugate momenta, we imagine this configuration as one mode in an orthogonal mode expansion of the conjugate momenta. Then the question is whether the presence of this one mode in a small (sub)volume has any effect on topological tunneling. As we now explain, we are easily able to produce this configuration of conjugate momenta.

To find the desired conjugate momentum, we begin with a continuum classical, instanton-like vector potential used by the Columbia group to investigate zero modes of the domain wall Dirac operator many years ago [6]. It takes the form

$$A_\mu(x) = -i \sum_{j=1}^3 \eta^{j\mu\nu} \lambda^j \frac{x_\nu}{x^2 + \rho(r)^2} , \quad \rho(r) = \rho_0 \left(1 - \frac{r}{r_{\max}}\right) \Theta(r_{\max} - r) \quad (1)$$

where A_μ is the gauge field potential, x_ν is the space-time coordinate, r is the magnitude of x , λ^j , $j = 1, 2, 3$ are the first three Gell-Mann matrices, $\eta^{j\mu\nu}$ is as in [7], Θ is the usual Heaviside function and ρ_0 is the instanton radius. Note that as $r_{\max} \rightarrow \infty$ this becomes a continuum instanton. With r_{\max} finite, outside of r_{\max} the configuration is strictly a gauge transformation. This continuum field is then exponentiated to create a lattice configuration of link matrices, $U_\mu(x)$. Finally, the lattice equivalent of the continuum transformation to singular gauge is applied:

$$U_\mu(x) \rightarrow g(x) U_\mu(x) g^{-1}(x + a_\mu) , \quad g(x) = \sum_{j=0}^3 \frac{x_j \lambda^j}{|x|} \quad (2)$$

where λ^0 is the identity matrix. Provided that the instanton center is not on a lattice site, this transformation is well defined everywhere. After the transformation, all links lying entirely outside r_{\max} are equal to the unit matrix so the configuration exactly obeys the usual periodic boundary conditions if r_{\max} is less than half the lattice size. We will refer to this field configuration as a cut-off instanton.

We want to find a configuration of conjugate momenta to use at the beginning of a molecular dynamics trajectory that will yield the above cut-off instanton at the end of the trajectory. To do that, we first create a cut-off instanton on a 16^4 lattice, with $\rho_0 = 4$ and $r_{\max} = 6$. Its center is at $(8.5, 8.5, 8.5, 8.5)$ and we denote its links by U . We make second instanton with the same center and r_{\max} , but with $\rho_0 = 3.9$, and denote its links by V . We then create a configuration of conjugate momenta by taking the traceless anti-hermitian part of $U \cdot V^{-1}$. We set the norm of this conjugate momentum mode to unity and then find that if we multiply it by 3, the instanton is essentially gone after 78 MD steps with step size = 0.025, or a total MD time of 1.95, as seen in Figure 1. Setting the gauge links to unity, reversing the conjugate momenta and restarting the MD evolves the lattice to similar topological charge as was there initially.

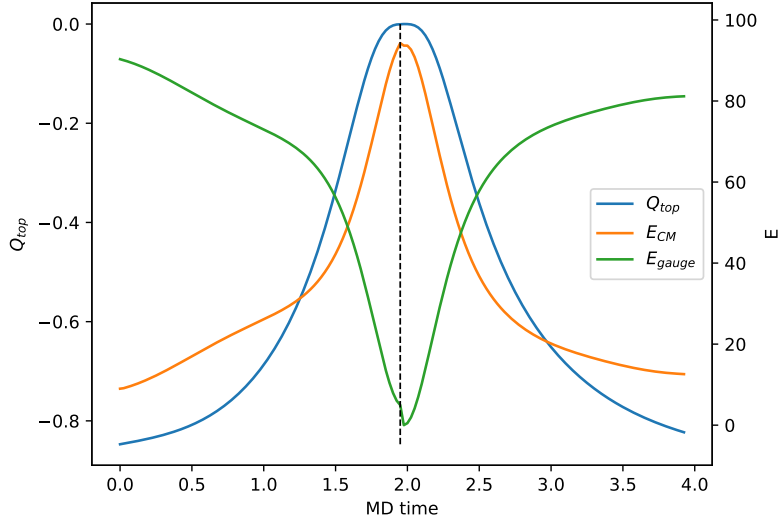


Figure 1: The molecular dynamics evolution of a cut-off instanton. The initial configuration is a cut-off instanton with conjugate momenta initialized as detailed in the text. Initially $Q_{\text{top}} = -0.847$ (simple plaquette definition, since this is a smooth field) and all of the action is in the gauge fields (green line). After 1.95 MD time units, Q_{top} is essentially zero and almost all of the action is in the conjugate momentum. At this point, all gauge links are set to unity, which produces the small discontinuity in the green line at about 2 MD time units, and the conjugate momenta are reversed. From the plot, one sees that the trajectory with reversed conjugate momentum produces a configuration with topology very similar to the initial configuration ($Q_{\text{top}} = -0.823$).

This simple classical evolution is an example of what we would like to try on a thermalized configuration in a small local volume set to a fixed gauge. Of course, the non-local evolution equations will distort whatever modes we may be trying to drive, making the realistic case far from this simple example. By having variable amplitude for the conjugate momentum modes, any given mode may persist farther in the evolution when it appears with a large amplitude. Before tackling the fixed gauge evolutions, we briefly revisit standard HMC for a pure gauge theory.

3. HMC tests with SU(3) Gauge Theory

As a starting point, we summarize the HMC and do basic tests for pure SU(3) gauge theory in four dimensions. At the beginning of a HMC trajectory, conjugate momenta $h_{j,\mu,a}$ are drawn from a Gaussian, to sample the distribution

$$\int [dh_{j,\mu,a}] \exp\left(-\sum_{j,\mu,a} h_{j,\mu,a}^2\right). \quad (3)$$

The conjugate momenta are used to form $H_{j,\mu} = \sum_a \lambda_a h_{j,\mu,a}$ and then the continuous time evolution equations are

$$\dot{U}_{j,\mu} = iH_{j,\mu} U_{j,\mu} \quad i\dot{H}_{j,\mu} = -\frac{\beta}{3} [U_{j,\mu} V_{j,\mu}]_{\text{TA}} \quad (4)$$

with $V_{j,\mu}$ the sum of staples and TA denoting the traceless, antihermitian part. For finite time steps and the standard leap-frog algorithm, this discretizes to

$$H_{j,\mu}(t + \Delta t/2) - H_{j,\mu}(t - \Delta t/2) = \dot{H}_{j,\mu}(t) \cdot \Delta t + O(\Delta t^3) \quad (5)$$

$$U_{j,\mu}(t + \Delta t) = \exp[iH(t + \Delta t/2) \cdot \Delta t] \cdot U_{j,\mu}(t) + O(\Delta t^3) \quad (6)$$

The left panel of Figure 2 shows the autocorrelation function for the plaquette for an evolution of 78k trajectories with the Wilson gauge action, $\beta = 6.0$ and a trajectory length of 1/2 MD time unit. By binning the data in large bins to get a good estimate of the standard deviation and comparing this with the naive standard deviation with no bins, one can get an estimate of the integrated autocorrelation time. In this case $\tau_{\text{INT}} = 10$ in units of trajectories. An estimate of the integrated autocorrelation time from the autocorrelation function, with a cutoff of 100 trajectories, gives $\tau_{\text{INT}}^{\text{cut}} \approx 7$. As one can see from the left panel of Figure 2, there is a secondary exponential tail in the autocorrelation function with $\tau_{\text{exp}} \sim 100$ and an amplitude of about 0.03, which makes up the difference between $\tau_{\text{INT}}^{\text{cut}}$ and τ_{INT} .

The appearance of such a long autocorrelation time in a lattice-local observable like the plaquette serves as a warning about the difficulty of measuring integrated autocorrelation times accurately. The right panel of Figure 2 shows the result for the same simulation with unit length trajectories. Here there is no evidence for the secondary exponential and the integrated autocorrelation time in trajectories drops from 10 to 5. However, in MD time units both simulations have τ_{INT} about 5, but with no long tail evidenced in the simulation with the longer trajectories.

We will use the ratio of the naive standard deviation and the standard deviation after blocking to determine integrated autocorrelation times for the simulations reported here. By comparing different bin sizes, we believe these estimates of the integrated autocorrelation times are accurate at the 10% level. Our main goal here is not a precise measurement of integrated autocorrelation times but we need some understanding of them to give us preliminary guidance during our explorations.

4. Variable Width Conjugate Momenta

We now consider changing the distribution of conjugate momenta from that given in Eq. 3 to

$$\int [dh_{j,\mu,a}] \exp\left(-\sum_{j,\mu,a} 2w_{j,\mu,a} h_{j,\mu,a}^2\right) \quad (7)$$

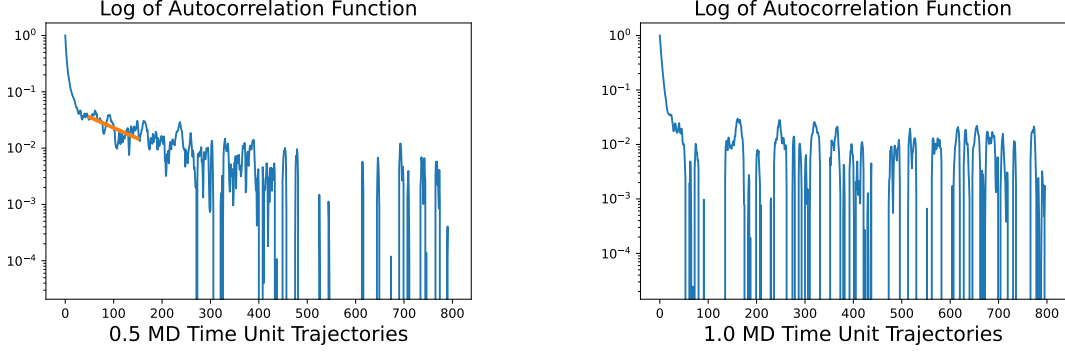


Figure 2: The autocorrelation function of the plaquette for an evolution of a 24^4 lattice with the Wilson gauge action and $\beta = 6.0$. In the left panel, the trajectory length is $1/2$ MD time unit with 78k total trajectories. For the right panel, the trajectory length is 1 MD time unit with 40k trajectories. Note the secondary decay (orange line) in the left panel. The integrated autocorrelation time, in MD time units, is the same for both simulations.

Variable CM width	Traj. length	Steps	dt	Total traj.	Acc.	IAC time traj.	IAC time MD	Plaquette
0	$1/2$	64	0.0078125	78k	0.81	10	5	0.593689(4)
0	1	160	0.00625	40k	0.90	5	5	0.593694(4)
0.1	$1/2$	64	0.0078125	10k	0.82	10-13	5-6.5	0.593683(12)
1.0	$1/2$	64	0.0078125	10k	0.79	8-10	4-5	0.593687(9)
4.0	$1/2$	100	0.005	11.8k	0.77	10	5	0.593686(10)

Table 1: Results for 24^4 lattices with $\beta = 6.0$. Values of δ_{CM} are given in the first column. The integrated autocorrelation times (IAC) are given in both trajectory and MD time units.

where we have introduced weights $w_{j,\mu,a}$ into the Gaussian integrals and $w_{j,\mu,a} = 1/2$ corresponds to the normal momentum draw. We choose to keep $\prod_{j,\mu,a} w_{j,\mu,a} = (1/2)^{32N^4}$. We choose the weights as $w = 1/2 + \delta_{\text{CM}} \cdot r$, where r is uniformly distributed between 0 and 1 and δ_{CM} is a constant that we can tune. For each conjugate momentum degree of freedom, we choose a w as above and then set the weight for another randomly chosen conjugate momentum to be $(4w)^{-1}$ to guarantee that the produce of the w values is $1/2$ per conjugate momentum degree of freedom.

Since we ultimately want to use variable width conjugate momenta that are non-local in space and color, we first want to see what effects variable width local conjugate momenta have. Table 1 has the results of 5 different simulations. One sees that the plaquette values all agree within errors and that the integrated autocorrelation times, in MD time units, are all the same, up to the accuracy that we know them. For $\delta_{\text{CM}} = 4$, one does have to reduce the MD step size noticeably to keep the acceptance similar. This is presumably due to the larger forces engendered by the larger values of some of the conjugate momentum when the width is variable.

5. Spatially Non-local Conjugate Momentum Modes

Consider an orthogonal transformation J from position space to another basis given by $\tilde{\phi} = J\phi$. We now draw \tilde{h} from $\exp(-\sum_i 2w_i \tilde{h}_i^2)$. Then in the original basis, we have

$$\int [dh_j] \exp\left(-\sum_{i,j} h_i K_{ij} h_j\right) \quad K_{ij} = \sum_m J_{im}^T (2w_m) J_{mk} \quad (8)$$

where the index can label lattice site and/or link and/or color. The Hamiltonian for the MD can then be written as

$$H = \frac{1}{2} \sum_{ij} h_i K_{ij} h_j - S_{\text{wil}}(U) \quad (9)$$

where $S_{\text{wil}}(U)$ is the Wilson gauge action. The non-locality of the conjugate momentum is represented by the matrix K_{ij} . This changes the initial conjugate momenta for an HMC trajectory to $h_i = \sum_j (J^{-1})_{i,j} r_j (2w_j)^{-1/2}$ where r is chosen from $\exp(-r^2)$. The only other change is in the conjugate momentum update, which now is given by

$$\dot{h}_i = (K^{-1})_{i,j} F_j. \quad (10)$$

where F_j is the standard HMC force for the j th degree of freedom. This allows one to see that h is updated by a non-local force related to K . Because this involves manipulations of the matrix K , we move to a smaller lattice volume where the inverse of K can easily be found. As mentioned earlier, if this method shows any promise, we would still restrict the non-local conjugate momentum modes to regions of order 0.5 fm in very large lattices.

Table 2 shows the results of simulations on a 4^4 lattice with $\beta = 5.6$. The first two rows are for standard HMC simulations, with a fixed width for the local conjugate momenta and two different trajectory lengths. Here one sees a modest improvement in integrated autocorrelation times (in MD units) for the longer trajectory length. For the last 3 rows, non-local conjugate momenta are used, with variable widths. For the 3rd and 4th rows, no gauge fixing is done, so the non-locality of the conjugate momenta should have no effect, except for perhaps more noise. No effects are visible, with the autocorrelation times and plaquette values similar to row 1.

The simulation in the last row of Table 2 has the lattice set to Landau gauge at the beginning of each trajectory. In this case, the non-local modes are not made effectively local, since the gauge is fixed. However, the lattice does not stay in Landau gauge during the evolution, so one has lost detailed balance in the HMC: the end point of a trajectory is not a point that would ever be the starting point for a subsequent trajectory. Evidence that this is a pathological case is clearly seen by the deviation of the plaquette value in row 5 from the other values in the table.

In order to see any effects from non-local conjugate momenta, we must be in a fixed gauge, which is maintained during the MD evolution. An easy gauge to try is maximal tree axial gauge, where all possible links are set to unity. (We don't fix the remaining global SU(3) symmetry, since we don't want to have to preserve this during the evolution.) The links that are gauge fixed do not change during the evolution and have no conjugate momentum. For all the links that do change, we create a single index listing them and then Fourier transform on that index to create the mapping matrix J mentioned above.

Variable CM width	Local	Traj. length	Steps	dt	Total traj.	Acc.	IAC time traj.	IAC time MD	Plaquette
0	yes	1/2	10	0.05	1M	0.81	48-52	24-26	0.53788(14)
0	yes	1	20	0.04	1M	0.90	17-19	17-19	0.53790(8)
0.3	no	1/2	10	0.05	100k	0.80	50	25	0.53790(46)
1.0	no	1/2	10	0.05	500k	0.78	42-46	21-23	0.53754(19)
1.0	no	1/2	10	0.05	11.8k	0.78	32-38	16-29	0.54116(16)

Table 2: 4^4 lattice with $\beta = 5.6$. The simulations described in the first 4 rows had no gauge fixing. The non-local and local results, for the same HMC parameters, are in good agreement. The simulation in the fifth row had Landau gauge fixing done at the start of each trajectory – a condition which is not preserved during the evolution – which makes the plaquette value incorrect.

Variable CM width	Local	Traj. length	Steps	dt	Total traj.	Acc.	IAC time traj.	IAC time MD	Plaquette
0	yes	1/2	10	0.05	1M	0.84	96-106	48-53	0.53820(20)
0	yes	1	20	0.05	1M	0.87	34-38	34-38	0.53802(12)
4.0	yes	1/2	20	0.025	5M	0.89	60-64	30-32	0.53790(7)
4.0	yes	1	40	0.025	5M	0.88	26-30	26-30	0.53779(5)
4.0	no	1	50	0.02	60k	0.92	36-38	18-19	0.53755(50)

Table 3: Axial gauge evolution of a 4^4 lattice with $\beta = 5.6$. The first four rows describe simulations with local conjugate momenta, while the last row is for non-local conjugate momenta. For this last case, the integrated autocorrelation time is substantially reduced.

The results of our axial gauge fixed evolutions are given in Table 3. The first 4 rows are results for local conjugate momenta and serve as a check of our code and simulations. Notice that for the first row, which is a standard HMC in axial gauge, the integrated autocorrelation times are 48-53 in MD time units, which is about twice the 24-26 values given in the first row of Table 2. As is generally expected, evolving in axial gauge is not very efficient, since variables that are spatially nearby may not influence each other easily because of all the fixed links in axial gauge. Rows 3 and 4 show some improvement in IAC by adding a variable width to the conjugate momenta. In row 5, we have variable width, non-local conjugate momenta and for this simulation we find the correct plaquette value and also a marked decrease in the integrated autocorrelation time from the standard HMC in axial gauge.

This last simulation provides evidence that non-local choices for conjugate momenta can compensate for the inefficiency inherent in an axial gauge HMC. Of course, the non-local axial gauge HMC is more numerically expensive than the standard HMC, due to the nonlocal character of the forces that enter the evolution, as seen in Eq. 10

6. Conclusions

We have presented our exploratory work on the effects of non-local conjugate momenta in simulations with the HMC algorithm. The driving idea behind this work is the question of whether,

in a large volume simulation, one might choose a number of small subvolumes (4^4 or 6^4 for current lattice spacings) and try to move the degrees of freedom in this subvolume in a coherent way through phase space. The results presented here show that if these subvolumes are fixed to axial gauge, which is preserved during a MD trajectory, non-local conjugate momenta can have an effect. In the case we studied, where the non-local conjugate momenta were not chosen with any particular insight, they compensate for the slower motion through phase space that comes from fixing to axial gauge. Whether more physically motivated non-local conjugate momenta, such as the choice discussed in Section 2 which creates an instanton from an initial lattice with all links equal to unity, offer further improvement is currently unknown, but can be investigated.

The author would like to acknowledge useful discussions with the Columbia lattice QCD group and the broader RBC-UKQCD collaboration. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC-0011941

References

- [1] S. Duane, A.D. Kennedy, B.J. Pendelton and D. Roweth, Phys. Lett. **B195** (1987) 216.
- [2] M.A. Clark and A.D. Kennedy, Phys. Rev. **D75** (2007) 011502.
- [3] S. Duane, R. Kenway, B.J. Pendelton and D. Roweth, Phys. Lett. **B176** (186) 143.
- [4] T. Nguyen, P. Boyle, N. Christ, Y.C. Jang and C. Jung, arXiv:2112.04556, <https://doi.org/10.48550/arXiv.2112.04556>.
- [5] S. Yamamoto, P. Boyle, T Izubuchi, L. Jun, C. Lehner and N. Matsumoto, arXiv:2502.05452, <https://doi.org/10.48550/arXiv.2502.05452>.
- [6] P. Chen, *et. al.* Phys. Rev. **D59** (1999) 054508.
- [7] G. 't Hooft, Phys. Rev. Lett **37** (1976) 8; Phys. Rev. **D14** (1976) 3432; Phys. Reports **142** (1986) 357.