Determining the strange quark mass for 2-flavour QCD∗


aInstitut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany
bSchool of Physics, University of Edinburgh, Edinburgh EH9 3JZ, UK
cDepartment of Mathematical Sciences, University of Liverpool, Liverpool L69 3BX, UK
dJohn von Neumann Institute NIC / DESY Zeuthen, D-15738 Zeuthen, Germany
eDeutsches Elektronen-Synchrotron DESY, D-22603 Hamburg, Germany
fKonrad-Zuse-Zentrum für Informationstechnik Berlin, D-14195 Berlin, Germany

Using the \(O(a)\) Symanzik improved action an estimate is given for the strange quark mass for unquenched \((n_f = 2)\) QCD. The determination is via the axial Ward identity (AWI) and includes a non-perturbative evaluation of the renormalisation constant. Numerical results have been obtained at several lattice spacings, enabling the continuum limit to be taken. Our results indicate a value for the strange quark mass (in the \(\overline{MS}\)-scheme at a scale of 2 GeV) in the range 100 – 130 MeV. A comparison is also made with other recent lattice determinations of the strange quark mass using dynamical sea quarks.

1. INTRODUCTION

Lattice methods allow, in principle, the complete ‘ab initio’ calculation of the fundamental parameters of QCD, such as quark masses. However quarks are not directly observable, being confined in hadrons and are thus not asymptotic states. So to determine their mass necessitates the use of a non-perturbative approach – such as lattice QCD or QCD sum rules. In this brief article, we report on our recent results for the strange quark mass for 2-flavour QCD in the \(\overline{MS}\)-scheme at a scale of 2 GeV, \(m_s^{\overline{MS}}(2\text{ GeV})\) (further details can be found in [1]).

The present phenomenological status is summarised by the Particle Data Group in [2] giving an estimate for the strange quark mass of 80 MeV < \(m_s^{\overline{MS}}(2\text{ GeV})\) < 130 MeV. This is a large band, and it is be hoped that lattice computations will reduce this significantly in the coming years.

∗Talk given by R. Horsley at the Workshop on Computational Hadron Physics, (Nicosia, Cyprus, September ’05).

2. RENORMALISATION GROUP INVARIANTS

Being confined, the mass of the quark, \(m_q^S(M)\), needs to be defined by giving a scheme, \(S\) and scale \(M\),

\[m_q^S(M) = Z_m^S(M)m_q^{\text{bare}},\]

(1)

and thus we need to find both the bare quark mass and the renormalisation constant. An added complication is that the \(\overline{MS}\)-scheme is a perturbative scheme, while more natural schemes which allow a non-perturbative definition of the renormalisation constants have to be used. It is thus convenient to first define a (non-unique) renormalisation group invariant (RGI) object, which is both scale and scheme independent by

\[m_q^{\text{RGI}} = \Delta Z_m^S(M)m_q^S(M) = Z_m^S m_q^{\text{bare}},\]

(2)

where we have

\[\Delta Z_m^S(M) = [2b_0g^S(M)^2]^{-\frac{dm_s}{dm}} \times\]
where it would appear that the expansion for \( M S \) has been determined we can then easily change scale using a higher scale is safer (which is chosen in practice). However, when the RGI quantity \( p \) is also choose to expand the \( S \) in terms of \( \gamma \). The \( S \) are known functions (with leading coefficients \( b_0, d_{m0} \)) and \( m \) respectively .

We have generated results for one, two, three- and four-loop results for \( m \), \( b \); other choices, of course, are also possible.

In Figs. 1 and 2 we show the results of solving eq. (3) as a function of the scale \( M \equiv \mu \) and \( M \equiv \mu \) for both the \( M S \) scheme using a higher scale is safer (which is chosen in practice). However, when the RGI quantity has been determined we can then easily change from one scheme to another. Of course these definitions also hold for other operators, see section 4.

The \( S \) and \( \gamma \) functions (with leading coefficients \( b_0, d_{m0} \)) are known perturbatively up to a certain order. In the \( M S \) scheme the first four coefficients are known, \( [3,4] \), and this is also true for the RI'-MOM scheme \( [5,6] \) (which is a suitable scheme for both perturbative and non-perturbative, NP, applications, see section 4). Note that in the RI'-MOM scheme we also choose to expand the \( S \) and \( \gamma \) functions in terms of \( g^{\overline{MS}} \); other choices, of course, are also possible.

In Figs. 1 and 2 we show the results of solv

\[
\exp \left\{ \int_0^{\infty} d\xi \left[ \frac{\gamma_\mu(\xi)}{\beta_0(\xi)} + \frac{d_{m0}}{b_0(\xi)} \right] \right\}.
\]

(3)

The \( S \) and \( \gamma \) functions (with leading coefficients \( b_0, d_{m0} \)) are known perturbatively up to a certain order. In the \( M S \) scheme the first four coefficients are known, \( [3,4] \), and this is also true for the RI'-MOM scheme \( [5,6] \) (which is a suitable scheme for both perturbative and non-perturbative, NP, applications, see section 4). Note that in the RI'-MOM scheme we also choose to expand the \( S \) and \( \gamma \) functions in terms of \( g^{\overline{MS}} \); other choices, of course, are also possible.

In Figs. 1 and 2 we show the results of solving eq. (3) as a function of the scale \( M \equiv \mu \) and \( M \equiv \mu \) for both the \( M S \) scheme using a higher scale is safer (which is chosen in practice). However, when the RGI quantity has been determined we can then easily change from one scheme to another. Of course these definitions also hold for other operators, see section 4.

\[
\exp \left\{ \int_0^{\infty} d\xi \left[ \frac{\gamma_\mu(\xi)}{\beta_0(\xi)} + \frac{d_{m0}}{b_0(\xi)} \right] \right\}.
\]

(3)

The \( S \) and \( \gamma \) functions (with leading coefficients \( b_0, d_{m0} \)) are known perturbatively up to a certain order. In the \( M S \) scheme the first four coefficients are known, \( [3,4] \), and this is also true for the RI'-MOM scheme \( [5,6] \) (which is a suitable scheme for both perturbative and non-perturbative, NP, applications, see section 4). Note that in the RI'-MOM scheme we also choose to expand the \( S \) and \( \gamma \) functions in terms of \( g^{\overline{MS}} \); other choices, of course, are also possible.

In Figs. 1 and 2 we show the results of solving eq. (3) as a function of the scale \( M \equiv \mu \) and \( M \equiv \mu \) for both the \( M S \) scheme using a higher scale is safer (which is chosen in practice). However, when the RGI quantity has been determined we can then easily change from one scheme to another. Of course these definitions also hold for other operators, see section 4.

\[
\exp \left\{ \int_0^{\infty} d\xi \left[ \frac{\gamma_\mu(\xi)}{\beta_0(\xi)} + \frac{d_{m0}}{b_0(\xi)} \right] \right\}.
\]

(3)

The \( S \) and \( \gamma \) functions (with leading coefficients \( b_0, d_{m0} \)) are known perturbatively up to a certain order. In the \( M S \) scheme the first four coefficients are known, \( [3,4] \), and this is also true for the RI'-MOM scheme \( [5,6] \) (which is a suitable scheme for both perturbative and non-perturbative, NP, applications, see section 4). Note that in the RI'-MOM scheme we also choose to expand the \( S \) and \( \gamma \) functions in terms of \( g^{\overline{MS}} \); other choices, of course, are also possible.

In Figs. 1 and 2 we show the results of solving eq. (3) as a function of the scale \( M \equiv \mu \) and \( M \equiv \mu \) for both the \( M S \) scheme using a higher scale is safer (which is chosen in practice). However, when the RGI quantity has been determined we can then easily change from one scheme to another. Of course these definitions also hold for other operators, see section 4.
where
\[ \frac{r_0 m_{ps}}{(r_0 m_{ps})^2} = c_a^{\text{REG}} + c_b^{\text{REG}} \left(r_0 m_{ps}\right)^2 + c_c^{\text{REG}} \left(r_0 m_{ps}\right)^2 + c_d^{\text{REG}} \left((r_0 m_{ps})^2 - 2(r_0 m_{ps})^2 \ln(r_0 m_{ps})^2 \right). \] 

\[ m_{ps}, m_{ps}^S \] are the valence and sea pseudoscalar masses respectively (both using mass degenerate quarks). The first term is the leading order, LO, result in \( \chi PT \) while the remaining terms come from the next non-leading order, NLO, in \( \chi PT \). Thus we see from eq. (5) that to NLO, we can first determine \( \beta \) defined for fixed \( \chi PT \) while the remaining terms come in terms of the hopping parameter by
\[ \frac{m}{2a} \left( \frac{1}{\kappa_q} - \frac{1}{\kappa_q^S} \right), \] 

or the axial Ward identity, AWI, which is the approach employed here. Imposing the AWI on the lattice for mass degenerate quarks, we have
\[ \partial_\mu A_\mu = 2\tilde{m}_q\not{\cal P} + O(a^2), \] 

and \( A \) and \( \not{\cal P} \) are the \( O(a) \) improved\(^3\) unrenormalised axial current and pseudoscalar density respectively and \( \tilde{m}_q \) is the AWI quark mass. So by forming two-point correlation functions with \( \not{\cal P} \) in the usual way, this bare quark mass can be determined
\[ a\tilde{m}_q \equiv \frac{\langle \partial_\mu A_\mu(t)\not{\cal P}(0) \rangle}{2\langle \not{\cal P}(t)\not{\cal P}(0) \rangle}. \] 

\(^2\) This is valid for both valence and sea quarks. \( \kappa_q^S \) is defined for fixed \( \beta \) by the vanishing of the pseudoscalar mass, i.e. \( m_{ps}(\kappa_q^S, \kappa_q^S) = 0 \). \( \kappa_q^c \) has been determined in [14].

\(^3\) The improvement term to the axial current, \( \partial_\mu \not{\cal P} \) together with improvement coefficient \( c_A \) has been included. The mass improvement terms, together with their associated difference in improvement coefficients, \( b_A, b_P \) appear to be small and have been ignored here.

We have found results for four \( \beta \)-values: 5.20, 5.25, 5.29, 5.40, each with several (three or more) sea quark masses and a variety of valence quark masses, \( [11] \).

Furthermore upon renormalisation we have
\[ A_\mu^R = Z_A A_\mu, \quad \not{\cal P}_{\beta}(M) = Z_{\beta}(M)\not{\cal P}, \] 
giving from eqs. (6) and (7)
\[ Z_{\beta}^{\text{REG}} = \Delta_{\text{REG}} Z_{\beta}(M) \frac{Z_A}{Z_{\beta}(M)}. \] 

As mentioned before, we use the RI’-MOM scheme, \( [5] \). This scheme considers amputated Green’s functions (practically in the Landau gauge) with an appropriate operator insertion, here either \( A \) or \( \not{\cal P} \). The renormalisation point is fixed at some momentum scale \( p^2 = \mu_p^2 \), and thus we have
\[ Z_{\beta}^{\text{RI’-MOM}}(\mu_p) = \frac{Z_{\beta}^{\text{RI’-MOM}}(p)}{\text{tr} \left[ \Gamma(p) \Gamma_{\text{Born}}^{-1}(p) \right]_{p^2=\mu_p^2}}, \] 

where \( \Gamma_{\text{O}} \) are one-particle irreducible (1PI) vertex functions, and \( Z_q \) is the wave-function renormalisation. (Our variation of the implementation of this method is described in \([12]\).) This determines \( D_{\text{RI’-MOM}} \) from which a chiral extrapolation, here using the sea quarks only, may be made to the chiral limit. For \( Z_{\beta} \) we make a linear extrapolation in \( am_q, Z_A = A_A + B_A am_q \), while for \( Z_{\beta}^{\text{RI’-MOM}} \) we must first subtract out a pole in the quark mass, \( [13] \), which occurs due to chiral symmetry breaking. We thus make a fit of the form
\[ (Z_{\beta}^{\text{RI’-MOM}})^{-1} = A_P + B_P/am_q. \] 

We now have all the components, namely \( A_A, A_P \) and \( \Delta Z_{\beta}^{\text{RI’-MOM}} \) necessary to compute \( Z_{\beta}^{\text{REG}} \) and hence \( r_0 m_{q,ps}^{\text{REG}} \). In Fig. 8 we show \( Z_{\beta}^{\text{REG}} \) for \( \beta = 5.20, 5.25, 5.29 \) and 5.40. These should be independent of the scale \( am_p \) at least for larger values. This seems to be the case, we make a phenomenological fit to account for residual effects.

5. COMPARISON OF \( Z_{\beta}^{\text{REG}} \) WITH OTHER METHODS

As many computations of the strange quark mass have used tadpole improved perturbation
theory together with a boosted coupling constant for the determination of the renormalisation constant, it is of interest to compare our results obtained in the previous section with this approach. Our variation of this method, tadpole-improved renormalisation-group-improved boosted perturbation theory or TRB-PT, is described in [14].

Regarding the lattice as a ‘scheme’, then from eq. (2) we can write

\[ m_{\text{RGI}} \approx \Delta Z_{\text{LAT}}(a) \tilde{m}(a) \]  

where the renormalisation-group-improved \( \Delta Z_{\text{LAT}}(a) \) is given by eq. (3). Furthermore in this ‘lattice’ scheme, we choose to use \( g_0^2 = g_0^2 / u_{0c}^4 \) where \( u_{0c}^4 = \langle \frac{1}{3} \text{Tr} U^2 \rangle \) (\( U^2 \) being the product of links around an elementary plaquette) rather than \( g_0 \), as series expansions in \( g_0^2 \) are believed to have better convergence. This is boosted perturbation theory. (We shall use chirally extrapolated plaquette values as determined in [1] at our \( \beta \)-values and so we add a subscript ‘c’ to \( u_0 \).) In the tadpole-improved method, noting that renormalisation constants for operators with no derivatives are \( \sim u_{0c} \), which indicates that \( Z_{\text{RGI}} u_{0c}^{-1} \) will converge faster then \( Z_{\text{RGI}} \) alone we re-write eq. (3) in the two loop approximation as

\[ Z_{\text{RGI}}^{(\text{TRB-PT})} \equiv \Delta Z_{\text{LAT}}(a) \]  

\[ = u_{0c} \left[ 2b_0 g_0^2 \left( 1 + \frac{p_1}{4} \frac{b_0}{b_1} \right) \right]^{q_1} \]  

where \( q_1 = (b_0 d_{m_1}^\text{LAT} - b_1 d_{m_0}) / (2 b_0 b_1) + (p_1 / 4) (b_0 / b_1) \) with \( p_1 = \frac{1}{4} \) being the first coefficient in the expansion of \( u_{0c} \). \( d_{m_1}^\text{LAT} \) may be found by relating the (known) perturbative result for \( Z_{\text{RGI}} \) to \( \Delta Z_{\text{LAT}} \).

In Fig. 3 we plot \( Z_{\text{RGI}}^{(\text{TRB-PT})} \) versus \( \beta \). Our NP results from section 4 are shown as filled circles. They are to be compared with the TRB-PT results denoted by empty squares. While there is a difference between the results, it is decreasing and thus may be primarily due to remnant \( O(a^2) \) effects, which disappear in the continuum limit. That various determinations of \( Z_{\text{RGI}}^{(\text{TRB-PT})} \) have different numerical values can be seen from the results

\[ \text{Fig. 3. } Z_{\text{RGI}}^{(\text{TRB-PT})} \text{ for } \beta = 5.20 \text{ (filled circles), } \beta = 5.25 \text{ (filled squares), } \beta = 5.29 \text{ (filled upper triangles), } \beta = 5.40 \text{ (filled lower triangles) together with fits } F(a_{\mu p}) = r_1 + r_2 (a_{\mu p})^2 + r_3 / (a_{\mu p})^2. \]
The difference between the two definitions is an O(\epsilon) effect. Using \(Z_{\text{TRB-PT}}^{\text{CON}}\) in \(Z_{\text{m}}^{\text{RGI}}\) leads, perhaps coincidentally, to very similar results to our NP results.

Investigating the possibility of O(\epsilon^2) differences a little further, we note that if we have two definitions of \(Z_{\text{m}}^{\text{RGI}}\) then if both are equally valid, forming the ratio should yield

\[
R_{\text{m}}^X = \frac{Z_{\text{m}}^{\text{RGI}}(X)}{Z_{\text{m}}^{\text{RGI}}} = 1 + O(\epsilon^2),
\]

where \(Z_{\text{m}}^{\text{RGI}}\) is the result of section 4 and \(X\) is some alternative definition (i.e. TRB-PT, ALPHA-Z\(A\), ALPHA-Z\(A^{\text{CON}}\)). In Fig. 4 we plot this ratio for these alternative definitions. The \(r_0/a\) values used for the \(x\)-axis are found by extrapolating the \(r_0/a\) results to the chiral limit. This extrapolation and results (for \((r_0/a)_{\text{c}}\)) are given in Fig. 4.

We see that (roughly) all three ratios extrapolate to 1 which implies that any of the four determinations of \(Z_{\text{m}}^{\text{RGI}}\) may be used. This includes the TRB-PT result. Of course other TI determinations might not have this property, and also their validity always has to be checked against a NP determination, so this result here is of limited use. It is also to be noted that different determinations can have rather different O(\epsilon^2) corrections, so a continuum extrapolation is always necessary.

### 6. RESULTS

Armed with \(Z_{\text{m}}^{\text{RGI}}\), we can now find \(r_0m_q^{\text{RGI}}\) and hence the ratio \(r_0m_q^{\text{RGI}}/(r_0m_{ps})^2\), using the values of \(r_0/a\) given in [7]. In Fig. 5 we plot this ratio (against \((r_0m_{ps})^2\)) for \(\beta = 5.29\). Using eq. (4) to eliminate \(c_{\alpha}^{\text{RGI}}\) in eq. (3) in favour of \(c_{\alpha}^{\text{RGI}}\) where

\[
c_{\alpha}^{\text{RGI}} = \frac{r_0m_q^{\text{RGI}}}{(r_0m_K^+)^2 + (r_0m_{K^0})^2 - (r_0m_{\pi^0})^2},
\]

(15)
gives \(r_0m_q^{\text{RGI}}\) directly\(^5\) to NLO in our fit function.

We have restricted the quark masses to lie in the range \((r_0m_{ps})^2 < 5\), which translates to \(m_{ps} \lesssim 880\) MeV, which is hopefully within the range of validity of low order \(\chi\)PT results. (Using \(r_0/a\), rather than their chirally extrapolated values for example, tends to give less variation in the ratio \(r_0m_q^{\text{RGI}}/(r_0m_{ps})^2\) so we expect LO \(\chi\)PT to be markedly dominant.) Varying this range

\(^5\)This is preferable to first determining \(c_{\alpha}^{\text{RGI}}\) and \(c_{\ell}^{\text{RGI}}\) i = b, c, d by using eq. (3) and then substituting in eq. (3) as the direct fit reduces the final error bar on \(r_0m_q^{\text{RGI}}\).
from 4 to 6 and higher gave some idea of possible systematic errors. Thus finally, for each β-value we have determined \( r_0 m_{\pi^0} \) and can now perform the last extrapolation to the continuum limit.

Our derivation so far, although needing a secondary quantity such as \( r_0 / a \) for a unit, depends only on lattice quantities. Only at the last stage, with our direct fit did we need to give a physical scale to this unit. A popular choice is \( r_0 = 0.5 \) fm. However there are some uncertainties in this value; our derivation using the nucleon gave \( r_0 = 0.467 \) fm and so to give some idea of scale uncertainties, we shall also consider this value. (The main change when changing the scale comes from the \( r_0 \)s in eq. 4, as \( m_{\pi^0} \propto r_0 \), while changes in \( \Delta Z_{\pi^0} \) are only logarithmic.)

Using the value for \( [\Delta Z_{\pi^0}(2 \text{ GeV})]^{-1} \) obtained in section 2 from Fig. 4 to convert \( m_{\pi^0} \) to \( m_{\pi^0}(2 \text{ GeV}) \) gives the results shown in Fig. 7. Also shown is an extrapolation to continuum limit. We finally obtain the result

\[
m_{\pi^0}(2 \text{ GeV}) = (16)
\]

\[
\begin{cases}
117(6)(4)(6) \text{ MeV} & \text{for } r_0 = 0.5 \text{ fm} \\
111(6)(4)(6) \text{ MeV} & \text{for } r_0 = 0.467 \text{ fm}
\end{cases}
\]

where the first error is statistical. The second error is systematic \( \sim 3 \) MeV estimated by varying the fit interval for \( (r_0 m_{\pi^0})^2 \). We take a further systematic error on these results as being covered by the different \( r_0 \) values of about \( \sim 6 \) MeV. This is to be compared to our previous result using the VWI [10], which gave results of 126(5) MeV, 119(5) MeV for \( r_0 = 0.5 \) fm and \( 0.467 \) fm respectively. These results and extrapolation are also shown in Fig. 7.

7. COMPARISONS

It is also useful to compare our results with the results from other groups. In Fig. 8 we show some results for \( n_f = 2 \) and \( n_f = 2 + 1 \) flavours (keeping the aspect ratio approximately the same as in Fig. 7). A variety of actions, renormalisations, units and scales have been used (so the results have been plotted in physical units using...
Figure 8. Results for $m_{\text{MS}}(2 \text{ GeV})$ versus $a^2 \text{ fm}^2$ using the AWI (upper plot) and VWI (lower plot) methods. The results are presented with the collaborations preferred units and scales. Circles (together with a linear continuum extrapolation) are from this work and [10]; diamonds from [10]; squares from [17]; up triangles from [18]; down triangles from [19]; left triangles from [20]; right triangles from [21]. NPR denotes non-perturbative renormalisation, while TB-PT denotes tadpole-improved boosted perturbation theory. [20,21] are for $n_f = 2 + 1$ flavours; the other results are all for $n_f = 2$ flavours.

the authors preferred values). In particular the HPQCD-MILC-UKQCD [21] and HPQCD [20] collaborations use improved staggered fermions. These fermions having a (remnant) chiral symmetry are in the same situation as overlap/domain wall fermions where there is no distinction between VWI and AWI quark masses; the bare quark mass in the Lagrangian simply needs to be renormalised.

As seen earlier in section 5 it is noticeable that the (tadpole improved) perturbative results lie lower than the non-perturbatively renormalised results. Also results with $a < 0.09 \text{ fm}$ (i.e. $a^2 < 0.008 \text{ fm}^2$) appear to be reasonably consistent with each other (this is more pronounced for the AWI results than for the VWI results). While results for $a < 0.09$ show some lattice discretisation effects, using results at larger lattice spacings seems to give a fairly constant extrapolation to the continuum limit. A similar effect has also been seen elsewhere, for example in the determination of $r_0 \Lambda_{\text{MS}}^3$ for $n_f = 0$ flavours, [17], where coarse lattices also show this characteristic flattening of the data.

Finally, we compare these numbers with results from the QCD sum rule approach. A recent review of results from this method is given in [22], citing as a final result $m_{\text{MS}}(2 \text{ GeV}) = 99(28) \text{ MeV}$. This covers the lattice results in Fig. 8.

8. CONCLUSIONS

In this article we have estimated the strange quark mass for 2-flavour QCD and found the result in eq. (16), using $O(a)$-improved clover fermions and taking into consideration non-perturbative (NP) renormalisation, the continuum extrapolation of the lattice results and the use of chiral perturbation theory. The NLO chiral perturbation theory yields a correction of about 5% to the LO result, and the relevant low energy constants are in rough agreement with the phenomenological values.
Figure 7. Results for $m_{s}^{\perp}(2\text{ GeV})$ (filled circles) versus the chirally extrapolated values of $(a/r_{0})^{2}$ (as given in [7]) together with a linear extrapolation to the continuum limit. For comparison, we also give our previous result using the VWI, [10] (open squares).

In conclusion, although there is a spread of results, it would seem that the unquenched strange quark mass determined here is not lighter than the quenched strange quark mass and lies in the range of 100 – 130 MeV.

ACKNOWLEDGEMENTS

The numerical calculations have been performed on the Hitachi SR8000 at LRZ (Munich), on the Cray T3E at EPCC (Edinburgh), on the Cray T3E at NIC (Jülich) and ZIB (Berlin), as well as on the APE1000 and Quadrics at DESY (Zeuthen). We thank all institutions. This work has been supported in part by the EU Integrated Infrastructure Initiative Hadron Physics (I3HP) under contract RII3-CT-2004-506078 and by the DFG under contract FOR 465 (Forschergruppe Gitter-Hadronen-Phänomenologie).

REFERENCES

1. QCDSF-UKQCD Collaboration: M. Göckeler et al., hep-lat/0601004.