Proof of the Folk Theorem

Here is a sketch of the proof of the folk theorem by Shoham, following Osborne and Rubinstein.

First, some notation and definitions.
Consider a game $G = (N, (A_i), (u_i))$. Let $v_i$ be the min-max value for player $i$, i.e.
$$v_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_{-i}, a_i)$$
Notice that in this definition players are not allowed to randomize. If we use the standard definition of $v_i$ that allows randomizations, the Folk theorem would still hold, though proofs would become more involved. (Intuitively, randomization isn’t needed in the repeated setting, since we can simulate frequencies of play). We say that a payoff profile $(r_i)$ is enforceable if $r_i \geq v_i$ for some $i$. We say that it is feasible if $r_i$ can be written as $\sum_{a \in A} \alpha_a u_i(a)$, where $A$ is the joint action space of $G$, for some $\alpha_a$’s that are rational, non-negative, and $\sum_{a \in A} \alpha_a = 1$. (i.e. $(r_i)$ is a convex rational combination of all outcomes in $G$).

Now we are ready to prove two parts of the theorem.

(The Folk Theorem: Part 1) For any game $G$ and any payoff pair $(r_i)$, if $(r_i)$ is a Nash equilibrium payoff profile of the average reward infinitely repeated game of $G$ then it is an enforceable payoff profile of $G$

Proof:
Suppose $(r_i)$ is not enforceable, i.e. $r_i < v_i$ for some $i$. Then consider a deviation of player $i$ to $b_i(s_{-i}(h))$ for any history $h$ of the repeated game, where $b_i$ is any best-response action in the one-shot game and $s_{-i}(h)$ is the (repeated) equilibrium strategy of other players. By definition of $b_i$, player $i$ would receive a payoff of at least $v_i$ in every stage game using this strategy. Thus, his payoff on average would also be at least $v_i > r_i$, and hence $(r_i)$ couldn’t be a Nash equilibrium, completing the proof.

(The Folk Theorem: Part 2) Suppose $(r_i)$ is a feasible enforceable payoff profile of $G$. Then it is a payoff profile of some equilibrium of the average reward infinitely repeated game of $G$.

Proof:
Suppose $(r_1, r_2)$ is a feasible enforceable payoff profile. Then we can write it as $r_i = \sum_{a \in A} (\frac{\beta_a}{\gamma}) u_i(a)$, where $\beta_a$ and $\gamma$ are non-negative integers. (Recall that $\alpha_a$ were required to be rational. So we can take $\gamma$ to be their common denominator). Since the combination was convex, we have $\gamma = \sum_{a \in A} \beta_a$. 

1
We’re going to construct a strategy profile that will cycle through all outcomes $a \in A$ of $G$ with cycles of length $\gamma$, each cycle repeating action $a$ exactly $\beta_a$ times. Let $(a^t)$ be such a sequence of outcomes.

Let’s define a strategy $s_i$ of player $i$ to be a grim (trigger) version of playing $(a^t)$: if nobody deviates, then $s_i$ plays $a^i_t$ in period $t$. However, if there was a period $t'$ in which some player $j \neq i$ deviated, then $s_i$ will play $(p_{-j})_i$, where $(p_{-j})$ is a solution to the minimization problem in the definition of $v_j$.

First note, that if everybody plays according to $s_i$, then, by construction, player $i$ receives average payoff of $r_i$ (look at averages over periods of length $\gamma$).

It is easy to see that this strategy profile is a Nash equilibrium: suppose everybody plays according to $s_i$, and player $j$ deviates at some point. Then, forever after, player $j$ will receive his min max payoff $v_j \leq r_j$, rendering the deviation unprofitable.